

=> d his full

(FILE 'HOME' ENTERED AT 15:05:06 ON 01 DEC 2005)

L1 FILE 'LREGISTRY' ENTERED AT 15:05:20 ON 01 DEC 2005
0 SEA ABB=ON NINDFDED|DEYVDN/SQSFP

L2 FILE 'REGISTRY' ENTERED AT 15:05:44 ON 01 DEC 2005
58109 SEA ABB=ON NINDFDED|DEYVDN/SQSFP
L3 27 SEA ABB=ON NINDFDED|DEYVDN/SQSP
L4 ANALYZE L3 1- LC : 4 TERMS
D

FILE 'REGISTRY' ENTERED AT 15:07:04 ON 01 DEC 2005
D QUE L3
D RN CN SQL KWIC NTE L3 1-27

L5 FILE 'CAPLUS, TOXCENTER, USPATFULL' ENTERED AT 15:07:30 ON 01 DEC 2005
43 SEA ABB=ON L3
L6 31 DUP REM L5 (12 DUPLICATES REMOVED)
ANSWERS '1-22' FROM FILE CAPLUS
ANSWERS '23-31' FROM FILE USPATFULL
D IBIB ED ABS HITRN 1-31

FILE 'HOME' ENTERED AT 15:08:17 ON 01 DEC 2005

FILE HOME

FILE LREGISTRY
LREGISTRY IS A STATIC LEARNING FILE

NEW CAS INFORMATION USE POLICIES, ENTER HELP USAGETERMS FOR DETAILS.

FILE REGISTRY
Property values tagged with IC are from the ZIC/VINITI data file
provided by InfoChem.

STRUCTURE FILE UPDATES: 29 NOV 2005 HIGHEST RN 868943-57-1
DICTIONARY FILE UPDATES: 29 NOV 2005 HIGHEST RN 868943-57-1

New CAS Information Use Policies, enter HELP USAGETERMS for details.

TSCA INFORMATION NOW CURRENT THROUGH JULY 14, 2005

Please note that search-term pricing does apply when
conducting SmartSELECT searches.

*
* The CA roles and document type information have been removed from *
* the IDE default display format and the ED field has been added, *
* effective March 20, 2005. A new display format, IDERL, is now *
* available and contains the CA role and document type information. *
*

Structure search iteration limits have been increased. See HELP SLIMITS
for details.

Checkal
JFR
12-6-2005

REGISTRY includes numerically searchable data for experimental and predicted properties as well as tags indicating availability of experimental property data in the original document. For information on property searching in REGISTRY, refer to:

<http://www.cas.org/ONLINE/UG/regprops.html>

FILE CAPLUS

Copyright of the articles to which records in this database refer is held by the publishers listed in the PUBLISHER (PB) field (available for records published or updated in Chemical Abstracts after December 26, 1996), unless otherwise indicated in the original publications. The CA Lexicon is the copyrighted intellectual property of the American Chemical Society and is provided to assist you in searching databases on STN. Any dissemination, distribution, copying, or storing of this information, without the prior written consent of CAS, is strictly prohibited.

FILE COVERS 1907 - 1 Dec 2005 VOL 143 ISS 23

FILE LAST UPDATED: 30 Nov 2005 (20051130/ED)

Effective October 17, 2005, revised CAS Information Use Policies apply. They are available for your review at:

<http://www.cas.org/infopolicy.html>

FILE TOXCENTER

FILE COVERS 1907 TO 29 Nov 2005 (20051129/ED)

This file contains CAS Registry Numbers for easy and accurate substance identification.

New CAS Information Use Policies, enter HELP USAGETERMS for details.

TOXCENTER has been enhanced with new files segments and search fields. See HELP CONTENT for more information.

TOXCENTER thesauri in the /CN, /CT, and /MN fields incorporate the MeSH 2005 vocabulary. See <http://www.nlm.nih.gov/mesh/> and http://www.nlm.nih.gov/pubs/techbull/nd04/nd04_mesh.html for a description of changes.

FILE USPATFULL

FILE COVERS 1971 TO PATENT PUBLICATION DATE: 1 Dec 2005 (20051201/PD)

FILE LAST UPDATED: 1 Dec 2005 (20051201/ED)

HIGHEST GRANTED PATENT NUMBER: US6971121

HIGHEST APPLICATION PUBLICATION NUMBER: US2005268363

CA INDEXING IS CURRENT THROUGH 1 Dec 2005 (20051201/UPCA)

ISSUE CLASS FIELDS (/INCL) CURRENT THROUGH: 1 Dec 2005 (20051201/PD)

REVISED CLASS FIELDS (/NCL) LAST RELOADED: Oct 2005

USPTO MANUAL OF CLASSIFICATIONS THESAURUS ISSUE DATE: Oct 2005

```
>>> USPAT2 is now available. USPATFULL contains full text of the      <<<
>>> original, i.e., the earliest published granted patents or         <<<
>>> applications. USPAT2 contains full text of the latest US         <<<
>>> publications, starting in 2001, for the inventions covered in     <<<
>>> USPATFULL. A USPATFULL record contains not only the original     <<<
```

>>> published document but also a list of any subsequent <<<
>>> publications. The publication number, patent kind code, and <<<
>>> publication date for all the US publications for an invention <<<
>>> are displayed in the PI (Patent Information) field of USPATFULL <<<
>>> records and may be searched in standard search fields, e.g., /PN, <<<
>>> /PK, etc. <<<

>>> USPATFULL and USPAT2 can be accessed and searched together <<<
>>> through the new cluster USPATALL. Type FILE USPATALL to <<<
>>> enter this cluster. <<<
>>> <<<
>>> Use USPATALL when searching terms such as patent assignees, <<<
>>> classifications, or claims, that may potentially change from <<<
>>> the earliest to the latest publication. <<<

This file contains CAS Registry Numbers for easy and accurate
substance identification.

=>

=> fil reg; d que 13
FILE 'REGISTRY' ENTERED AT 15:07:04 ON 01 DEC 2005
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2005 American Chemical Society (ACS)

Property values tagged with IC are from the ZIC/VINITI data file
provided by InfoChem.

STRUCTURE FILE UPDATES: 29 NOV 2005 HIGHEST RN 868943-57-1
DICTIONARY FILE UPDATES: 29 NOV 2005 HIGHEST RN 868943-57-1

New CAS Information Use Policies, enter HELP USAGETERMS for details.

TSCA INFORMATION NOW CURRENT THROUGH JULY 14, 2005

Please note that search-term pricing does apply when
conducting SmartSELECT searches.

*
* The CA roles and document type information have been removed from *
* the IDE default display format and the ED field has been added, *
* effective March 20, 2005. A new display format, IDERL, is now *
* available and contains the CA role and document type information. *
*

Structure search iteration limits have been increased. See HELP SLIMITS
for details.

REGISTRY includes numerically searchable data for experimental and
predicted properties as well as tags indicating availability of
experimental property data in the original document. For information
on property searching in REGISTRY, refer to:

<http://www.cas.org/ONLINE/UG/regprops.html>

L3 27 SEA FILE=REGISTRY ABB=ON NINDFDED|DEYVDN/SQSP

=> d rn cn sql kwic nte 13 1-27; fil capl toxcenter uspatf; s 13

L3 ANSWER 1 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN
RN 859621-51-5 REGISTRY
CN L-Valine, L-methionyl-L-arginyl-L-valyl-L-lysyl-L-threonyl-L-phenylalanyl-
L-valyl-L-isoleucyl-L-leucyl-L-cysteinyl-L-cysteinyl-L-alanyl-L-leucyl-L-
glutaminyl-L-tyrosyl-L-valyl-L-alanyl-L-tyrosyl-L-threonyl-L-asparaginy-L-
alanyl-L-asparaginy-L-isoleucyl-L-asparaginy-L- α -aspartyl-L-
phenylalanyl-L- α -aspartyl-L- α -glutamyl-L- α -aspartyl-L-
tyrosyl-L-phenylalanylglycyl-L-seryl-L- α -aspartyl- (9CI) (CA INDEX
NAME)

OTHER NAMES:

CN 4: PN: WO2005068495 SEQID: 4 claimed protein
CN Fibroin (Bombyx mori H-chain N-terminal fragment)
SQL 35

SEQ 1 MRVKTFVILC CALQYVAYTN ANINDFDEDY FGSDV

=====

HITS AT: 22-29

L3 ANSWER 2 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN

RN 815501-68-9 REGISTRY

CN Protein (Staphylococcus epidermidis strain RP62A 383-amino acid) (9CI)
(CA INDEX NAME)

OTHER NAMES:

CN GenBank AAW54077

CN GenBank AAW54077 (Translated from: GenBank CP000029)

SQL 383

SEQ 201 TRRQFNRNAQ QQDSYNGITD NQPDEDTSSD QLYSDEYVDN EDKYSQFPKR

=====

HITS AT: 235-240

RELATED SEQUENCES AVAILABLE WITH SEQLINK

L3 ANSWER 3 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN

RN 803823-77-0 REGISTRY

CN 11: PN: JP2004339189 PAGE: 9 unclaimed sequence (9CI) (CA INDEX NAME)

SQL 120

SEQ 1 MRVTAFVILC CALQYATANN LHHHDEYVDN HGQLVERFTT RKHYERNAAT

=====

HITS AT: 25-30

L3 ANSWER 4 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN

RN 803823-75-8 REGISTRY

CN 1: PN: JP2004339189 PAGE: 8 unclaimed sequence (9CI) (CA INDEX NAME)

SQL 151

SEQ 1 MRVKTFVILC CALQYVAYTN ANINDFDEDY FGSDVTVQSS NTTDEIIRDA

=====

HITS AT: 22-29

L3 ANSWER 5 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN

RN 775416-81-4 REGISTRY

CN Protein (Staphylococcus aureus clone WO2002086097-SEQID-5635) (9CI) (CA
INDEX NAME)

OTHER NAMES:

CN 4656: PN: WO2002086097 SEQID: 5635 claimed protein

SQL 2368

SEQ 1501 AAADKKTQIE QTPNASQQEI NDAKQEVDTL LNQAKTNIDQ SSTDEYVDNA

=====

HITS AT: 1544-1549

RELATED SEQUENCES AVAILABLE WITH SEQLINK

L3 ANSWER 6 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN

RN 714954-21-9 REGISTRY

CN L-Aspartic acid, L-asparaginyL-L-isoleucyl-L-asparaginyL-L- α -
aspartyl-L-phenylalanyl-L- α -aspartyl-L- α -glutamyl- (9CI) (CA
INDEX NAME)

SQL 8

SEQ 1 NINDFDED

=====

HITS AT: 1-8

L3 ANSWER 7 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN
RN 672991-01-4 REGISTRY
CN Transcription-associated protein (Glycine max clone
PAT_MRT3847_89755C.1.pep fragment) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 2345: PN: US20040031072 SEQID: 274345 claimed protein
SQL 148

SEQ 51 EGGSKLDEYV DNCGPVTKSR DNIGEEMLLS HRSKEPGRNE LGDPLSTFAA
=====

HITS AT: 57-62

L3 ANSWER 8 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN
RN 660056-86-0 REGISTRY
CN Protein (Streptococcus pneumoniae strain 14453 clone US6699703-SEQID-4900
open reading frame-encoded) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 4900: PN: US6699703 SEQID: 4900 claimed protein
SQL 126

SEQ 51 PPLKVMALLV HGALQQYEHG YSLEDVYDLY DEYVDNNGDQ TTFMTEVLMP
=====

HITS AT: 81-86

L3 ANSWER 9 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN
RN 604934-18-1 REGISTRY
CN Protein (Staphylococcus epidermidis strain ATCC12228 gene SE0801) (9CI)
(CA INDEX NAME)

OTHER NAMES:

CN GenBank AAO04398
CN GenBank AAO04398 (Translated from: GenBank AE016746)
SQL 383

SEQ 201 TRRQFNRNAQ QQDSYNGITD NQPDEDTSSD QLYSDEYVDN EDKYSQFPKR
=====

HITS AT: 235-240

RELATED SEQUENCES AVAILABLE WITH SEQLINK

L3 ANSWER 10 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN
RN 483169-91-1 REGISTRY
CN GenBank CAA23432 (9CI) (CA INDEX NAME)

OTHER NAMES:

CN GenBank CAA23432 (Translated from: GenBank V00094)
SQL 168

SEQ 1 MRVKTFVILV CALQYVAYTN ANINDFDEDY FGSDVTVQSS NTTDEIIRDA
=====

HITS AT: 22-29

L3 ANSWER 11 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN
RN 482997-73-9 REGISTRY
CN GenBank AAF78030 (9CI) (CA INDEX NAME)

OTHER NAMES:

CN GenBank AAF78030 (Translated from: GenBank AF242774)
SQL 30

SEQ 1 MRVIAFVILC CALQYATAKN LRHHDEYVDN

HITS AT: 25-30

RELATED SEQUENCES AVAILABLE WITH SEQLINK

L3 ANSWER 12 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN
RN 482255-67-4 REGISTRY
CN GenBank AAA27838 (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AAA27838 (Translated from: GenBank M24222)
SQL 178

SEQ 1 MRVKTFVILC CALQYVAYTN ANINDFDEDY FGSDVTVQSS NTTDEIIRDA

HITS AT: 22-29

L3 ANSWER 13 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN
RN 482246-94-6 REGISTRY
CN GenBank CAA27612 (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank CAA27612 (Translated from: GenBank X03973)
SQL 178

SEQ 1 MRVKTFVILV CALQYVAYTN ANINDFDEDY FGSDVTVQSS NTTDEIIRDA

HITS AT: 22-29

L3 ANSWER 14 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN
RN 469866-34-0 REGISTRY
CN L-Asparagine, L-methionyl-L-arginyl-L-valyl-L-isoleucyl-L-alanyl-L-phenylalanyl-L-valyl-L-isoleucyl-L-leucyl-L-cysteinyl-L-cysteinyl-L-alanyl-L-leucyl-L-glutamyl-L-tyrosyl-L-alanyl-L-threonyl-L-alanyl-L-lysyl-L-asparaginyL-L-leucyl-L-arginyl-L-histidyl-L-histidyl-L- α -aspartyl-L- α -glutamyl-L-tyrosyl-L-valyl-L- α -aspartyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN Fibroin (Antheraea pernyi)
SQL 30

SEQ 1 MRVIAFVILC CALQYATAKN LRHHDEYVDN

HITS AT: 25-30

RELATED SEQUENCES AVAILABLE WITH SEQLINK

L3 ANSWER 15 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN
RN 467525-63-9 REGISTRY
CN Protein (Plasmodium falciparum strain 3D7 clone MAL4P2 gene PFD0380c) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank CAB62854
CN GenBank CAB62854 (Translated from: GenBank AL035475)
SQL 1629

SEQ 1151 MMVGTKDKKK NKKKKKKKKNK NKNYNNNNNN NKILEDDEYV DNIYYNNTNN

HITS AT: 1187-1192

L3 ANSWER 16 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN
RN 465605-62-3 REGISTRY

CN Protein (Plasmodium falciparum strain 3D7 gene PF14-0556) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN GenBank AAN37169

CN GenBank AAN37169 (Translated from: GenBank AE014825)

SQL 1338

SEQ 701 FNMNRNLPTF ADTLIIDEYV DNYWSENKLN NIDFRLFLQS WKVLNDCISF

==== ==

HITS AT: 717-722

L3 ANSWER 17 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN

RN 465598-80-5 REGISTRY

CN Protein (Plasmodium falciparum strain 3D7 gene PFB0440c) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN GenBank AAC71877

CN GenBank AAC71877 (Translated from: GenBank AE001395)

SQL 587

SEQ 501 MSSRLREYEI LDDEYVDNIE CLNKYVSVLN TNDVNIMDDR ERECSDDYSDE

=====

HITS AT: 513-518

L3 ANSWER 18 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN

RN 445314-07-8 REGISTRY

CN Antigen (Staphylococcus aureus clone ORF1381) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 566: PN: WO02059148 SEQID: 576 claimed protein

SQL 383

SEQ 201 TRRQFNRNAQ QQDSYNGITD NQPDEDTSSD QLYSDEYVDN EDKYSQFPKR

=====

HITS AT: 235-240

RELATED SEQUENCES AVAILABLE WITH SEQLINK

L3 ANSWER 19 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN

RN 437954-61-5 REGISTRY

CN Essential protein (Staphylococcus aureus clone WO0170955-SEQID-12389) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 2355: PN: WO0170955 SEQID: 12389 claimed protein

SQL 2368

SEQ 1501 AAADKKTQIE QTPNASQQEI NDAKQEVDTN LNQAKTNIDQ SSTDEYVDNA

=====

HITS AT: 1544-1549

RELATED SEQUENCES AVAILABLE WITH SEQLINK

L3 ANSWER 20 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN

RN 433270-43-0 REGISTRY

CN DNA (Staphylococcus aureus clone WO0170955-SEQID-8291 proliferation-associated gene) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 139: PN: WO0170955 SEQID: 8291 claimed protein

SQL 2368

SEQ 1501 AAADKKTQIE QTPNASQQEI NDAKQEVDTN LNQAKTNIDQ SSTDEYVDNA

=====

HITS AT: 1544-1549

****RELATED SEQUENCES AVAILABLE WITH SEQLINK****

L3 ANSWER 21 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN
RN 421018-05-5 REGISTRY
CN Protein (Staphylococcus epidermidis strain 19804 clone
US6380370-SEQID-3437 fragment) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 3094: PN: US6380370 SEQID: 3437 claimed protein
SQL 384

SEQ 201 NTRRQFNRNA QQQDSYNGIT DNQPDDETSS DQLYSDEYVD NEDKYSQFPK

=====

HITS AT: 236-241

L3 ANSWER 22 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN
RN 404318-03-2 REGISTRY
CN Fibroin (Antheraea yamamai) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN GenBank AAK83145
CN GenBank AAK83145 (Translated from: GenBank AF325500)
SQL 2655

SEQ 1 MRVTAFVILC CALQYATANN LHHHDEYVDN HGQLVERFTT RKHYERNAAT

=====

HITS AT: 25-30

L3 ANSWER 23 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN
RN 364143-92-0 REGISTRY
CN Protein (Staphylococcus aureus clone SAU102284 proliferation-associated
fragment) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 4085: PN: WO0170955 SEQID: 5635 claimed protein
SQL 2368

SEQ 1501 AAADKKTQIE QTPNASQQEI NDAKQEVDTL LNQAKTNIDQ SSTDEYVDNA

=====

HITS AT: 1544-1549

****RELATED SEQUENCES AVAILABLE WITH SEQLINK****

L3 ANSWER 24 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN
RN 341040-34-4 REGISTRY
CN Protein (Streptococcus epidermidis clone contig_0755_pos_5604_4453) (9CI)
(CA INDEX NAME)

OTHER NAMES:

CN 1026: PN: WO0134809 SEQID: 2426 claimed protein
SQL 383

SEQ 201 TRRQFNRNAQ QQDSYNGITD NQPDDETSSD QLYSDEYVDN EDKYSQFPKR

=====

HITS AT: 235-240

****RELATED SEQUENCES AVAILABLE WITH SEQLINK****

L3 ANSWER 25 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN
RN 336885-96-2 REGISTRY
CN Fibroin (Antheraea pernyi clone AP2) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN GenBank AAC32606
CN GenBank AAC32606 (Translated from: GenBank AF083334)
SQL 2639

SEQ 1 MRVIAFVILC CALQYATAKN LRHHDEYVDN HGQLVERFTT RKHFERNAAT

=====

HITS AT: 25-30

L3 ANSWER 26 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN
RN 303229-60-9 REGISTRY
CN Fibroin (silkworm strain p50 heavy chain) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN Fibroin (Bombyx mori strain p50 gene fib-H heavy chain)
CN GenBank AAF76983
CN GenBank AAF76983 (Translated from: GenBank AF226688)
SQL 5263

SEQ 1 MRVKTFVILC CALQYVAYTN ANINDFDEY FGSDVTQSS NTTDEIIRDA

=====

HITS AT: 22-29

L3 ANSWER 27 OF 27 REGISTRY COPYRIGHT 2005 ACS on STN
RN 257896-67-6 REGISTRY
CN Chromatinic RING finger protein, DRING ortholog (Plasmodium falciparum gene PFB0440c) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 17: PN: WO0025728 SEQID: 87 claimed protein
CN Protein (Plasmodium falciparum clone p3D7 chromosome 2 gene PFB0440)
CN RING finger-containing protein (Plasmodium falciparum clone 3D7 gene PFB0440c)

SQL 568

SEQ 451 SSSDSSNSNQ NNYINFMYNK KGKDIIVPMT KMSSRLREYE ILDDEYVDNI

=====

HITS AT: 494-499

FILE 'CAPLUS' ENTERED AT 15:07:30 ON 01 DEC 2005
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2005 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'TOXCENTER' ENTERED AT 15:07:30 ON 01 DEC 2005
COPYRIGHT (C) 2005 ACS

FILE 'USPATFULL' ENTERED AT 15:07:30 ON 01 DEC 2005
CA INDEXING COPYRIGHT (C) 2005 AMERICAN CHEMICAL SOCIETY (ACS)

L5 43 L3

=> dup rem 15

PROCESSING COMPLETED FOR L5

L6 31 DUP REM L5 (12 DUPLICATES REMOVED)
ANSWERS '1-22' FROM FILE CAPLUS
ANSWERS '23-31' FROM FILE USPATFULL

=> d ibib ed abs hitrn 1-31; fil hom

ACCESSION NUMBER: 2005:326915 CAPLUS Full-text
 DOCUMENT NUMBER: 142:330659
 TITLE: Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant *Staphylococcus aureus* strain and a biofilm-producing methicillin-resistant *Streptococcus epidermidis* strain
 AUTHOR(S): Gill, Steven R.; Fouts, Derrick E.; Archer, Gordon L.; Mongodin, Emmanuel F.; DeBoy, Robert T.; Ravel, Jacques; Paulsen, Ian T.; Kolonay, James F.; Brinkac, Lauren; Beanan, Mauren; Dodson, Robert J.; Daugherty, Sean C.; Madupu, Ramana; Angiuoli, Samuel V.; Durkin, A. Scott; Haft, Daniel H.; Vamathevan, Jessica; Khouri, Hoda; Utterback, Terry; Lee, Chris; Dimitrov, George; Jiang, Lingxia; Qin, Haiying; Weidman, Jan; Tran, Kevin; Kang, Kathy; Hance, Ioana R.; Nelson, Karen E.; Fraser, Claire M.
 CORPORATE SOURCE: The Institute for Genomic Research, Rockville, MD, USA
 SOURCE: Journal of Bacteriology (2005), 187(7), 2426-2438
 CODEN: JOBAAY; ISSN: 0021-9193
 PUBLISHER: American Society for Microbiology
 DOCUMENT TYPE: Journal
 LANGUAGE: English
 ED Entered STN: 18 Apr 2005

AB *Staphylococcus aureus* is an opportunistic pathogen and the major causative agent of numerous hospital- and community-acquired infections. *Staphylococcus epidermidis* has emerged as a causative agent of infections often associated with implanted medical devices. The .apprx.2.8-Mb genome of *S. aureus* COL, an early methicillin-resistant isolate, and the .apprx.2.6-Mb genome of *S. epidermidis* RP62a, a methicillin-resistant biofilm isolate, were sequenced. Comparative anal. of these and other staphylococcal genomes was used to explore the evolution of virulence and resistance between these two species. The *S. aureus* and *S. epidermidis* genomes are syntenic throughout their lengths and share a core set of 1681 open reading frames. Genome islands in nonsyntenic regions are the primary source of variations in pathogenicity and resistance. Gene transfer between staphylococci and low-GC-content gram-pos. bacteria appears to have shaped their virulence and resistance profiles. Integrated plasmids in *S. epidermidis* carry genes encoding resistance to cadmium and species-specific LPXTG surface proteins. A novel genome island encodes multiple phenol-soluble modulins, a potential *S. epidermidis* virulence factor. *S. epidermidis* contains the cap operon, encoding the polyglutamate capsule, a major virulence factor in *Bacillus anthracis*. Addnl. phenotypic differences are likely the result of single nucleotide polymorphisms, which are most numerous in cell envelope proteins. Overall differences in pathogenicity can be attributed to genome islands in *S. aureus* which encode enterotoxins, exotoxins, leukocidins, and leukotoxins not found in *S. epidermidis*.

IT 815501-68-9

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
 (amino acid sequence; evolution of virulence and resistance based on complete genome anal. of methicillin-resistant *Staphylococcus aureus* strain and biofilm-producing methicillin-resistant *S. epidermidis* strain)

REFERENCE COUNT: 54 THERE ARE 54 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

ACCESSION NUMBER: 2004:176539 CAPLUS Full-text
 DOCUMENT NUMBER: 140:176343
 TITLE: Nucleic acid and amino acid sequences relating to
 Streptococcus pneumoniae for diagnostics and
 therapeutics
 INVENTOR(S): Doucette-stamm, Lynn; Bush, David; Zeng, Qiandong;
 Opperman, Timothy; Houseweart, Chad Eric
 PATENT ASSIGNEE(S): Genome Therapeutics Corporation, USA
 SOURCE: U.S., 301 pp., Cont.-in-part of U.S. Ser. No. 107,433.
 CODEN: USXXAM
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 1
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
-----	---	----	-----	-----
US 6699703	B1	20040302	US 2000-583110	20000526
US 6800744	B1	20041005	US 1998-107433	19980630
US 2005136404	A1	20050623	US 2003-617320	20030710
PRIORITY APPLN. INFO.:			US 1997-51553P	P 19970702
			US 1998-85131P	P 19980512
			US 1998-107433	A2 19980630

ED Entered STN: 04 Mar 2004

AB The invention provides isolated polypeptide and nucleic acid sequences derived
 from Streptococcus pneumoniae that are useful in diagnosis and therapy of
 pathol. conditions. Thus, 2661 genomic DNA sequences are provided from S.
 pneumoniae strain 14453 and analyzed for the presence of open reading frames
 comprising at least 180 nucleotides and the start codons. Antibodies against
 the polypeptides, and methods for the production of the polypeptides are
 provided, as well as methods for the detection, prevention and treatment of
 pathol. conditions resulting from bacterial infection.

IT 660056-86-0

RL: BSU (Biological study, unclassified); DGN (Diagnostic use); PRP
 (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)
 (amino acid sequence; nucleic acid and amino acid sequences relating to
 Streptococcus pneumoniae for diagnostics and therapeutics)

REFERENCE COUNT: 14 THERE ARE 14 CITED REFERENCES AVAILABLE FOR THIS
 RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L6 ANSWER 3 OF 31 CAPLUS COPYRIGHT 2005 ACS on STN DUPLICATE 3

ACCESSION NUMBER: 2003:776159 CAPLUS Full-text
 DOCUMENT NUMBER: 139:271815
 TITLE: Genome-based analysis of virulence genes in a
 non-biofilm-forming Staphylococcus epidermidis strain
 (ATCC 12228)
 AUTHOR(S): Zhang, Yue-Qing; Ren, Shuang-Xi; Li, Hua-Lin; Wang,
 Yong-Xiang; Fu, Gang; Yang, Jian; Qin, Zhi-Qiang;
 Miao, You-Gang; Wang, Wen-Yi; Chen, Run-Sheng; Shen,
 Yan; Chen, Zhu; Yuan, Zheng-Hong; Zhao, Guo-Ping; Qu,
 Di; Danchin, Antoine; Wen, Yu-Mei
 CORPORATE SOURCE: Laboratory of Medical Molecular Virology, Shanghai
 Medical College, Fudan University, Shanghai, 200032,
 Peop. Rep. China
 SOURCE: Molecular Microbiology (2003), 49(6), 1577-1593
 CODEN: MOMIEE; ISSN: 0950-382X
 PUBLISHER: Blackwell Publishing Ltd.
 DOCUMENT TYPE: Journal
 LANGUAGE: English
 ED Entered STN: 03 Oct 2003

AB Staphylococcus epidermidis strains are diverse in their pathogenicity; some are invasive and cause serious nosocomial infections, whereas others are non-pathogenic commensal organisms. To analyze the implications of different virulence factors in Staphylococcus epidermidis infections, the complete genome of Staphylococcus epidermidis strain ATCC 12228, a non-biofilm forming, non-infection associated strain used for detection of residual antibiotics in food products, was sequenced. This strain showed low virulence by mouse and rat exptl. infections. The genome consists of a single 2,499,279 bp chromosome and 6 plasmids. The chromosomal G + C content is 32.1% and 2419 protein coding sequences (CDS) are predicted, among which 230 are putative novel genes. Compared to the virulence factors in Staphylococcus aureus, aside from δ -hemolysin and β -hemolysin, other toxin genes were not found. In contrast, the majority of adhesin genes are intact in ATCC 12228. Most strikingly, the ica operon coding for the enzymes synthesizing interbacterial cellular polysaccharide is missing in ATCC 12228 and rearrangements of adjacent genes are shown. No mec genes, IS256, IS257, were found in ATCC 12228. It is suggested that the absence of the ica operon is a genetic marker in commensal Staphylococcus epidermidis strains which are less likely to become invasive.

IT 604934-18-1

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)

(amino acid sequence; genome-based anal. of virulence genes in a non-biofilm-forming Staphylococcus epidermidis strain (ATCC 12228))

REFERENCE COUNT: 69 THERE ARE 69 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L6 ANSWER 4 OF 31 CAPLUS COPYRIGHT 2005 ACS on STN DUPLICATE 4

ACCESSION NUMBER: 2004:902756 CAPLUS Full-text
Correction of: 2002:832949

DOCUMENT NUMBER: 141:343454
Correction of: 137:346147

TITLE: Methods for identifying the target of a compound which inhibits cellular proliferation

INVENTOR(S): Carr, Grant J.; Xu, Howard H.; Foulkes, Gordon J.; Zamudio, Carlos; Haselbeck, Robert; Ohlsen, Kari L.; Zyskind, Judith W.; Wall, Daniel; Trawick, John D.; Yamamoto, Robert T.; Roemer, Terry; Jiang, Bo; Boone, Charles; Bussey, Howard

PATENT ASSIGNEE(S): Elitra Pharmaceuticals, Inc., USA

SOURCE: PCT Int. Appl., 640 pp.
CODEN: PIXXD2

DOCUMENT TYPE: Patent
LANGUAGE: English

FAMILY ACC. NUM. COUNT: 3

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2002086097	A2	20021031	WO 2002-US3987	20020208
WO 2002086097	A3	20030306		
WO 2002086097	C1	20041125		
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW			
RW:	GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB,			

GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, ML, MR, NE, SN, TD, TG

CA 2436216 AA 20021031 CA 2002-2436216 20020208
WO 2002086097 A2 20021031 WO 2002-XA3987 20020208

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN,
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH,
PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ,
UA, UG, US, UZ, VN, YU, ZA, ZM, ZW

RW: GH, GM, KE, LK, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AM, AZ, BY,
KG, KZ, MD, RU, TJ, TM, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB,
GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, ML, MR, NE, SN, TD, TG

WO 2002086097 A2 20021031 WO 2002-XB3987 20020208

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN,
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH,
PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ,
UA, UG, US, UZ, VN, YU, ZA, ZM, ZW

RW: GH, GM, KE, LK, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AM, AZ, BY,
KG, KZ, MD, RU, TJ, TM, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB,
GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, ML, MR, NE, SN, TD, TG

EP 1360335 A2 20031112 EP 2002-728338 20020208

R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT,
IE, SI, LT, LV, FI, RO, MK, CY, AL, TR

JP 2004528846 T2 20040924 JP 2002-583612 20020208

PRIORITY APPLN. INFO.: US 2001-267636P P 20010209
WO 2002-US3987 W 20020208

ED Entered STN: 29 Oct 2004

AB The invention relates to cultures or collections of strains which overexpress or underexpress gene products required for the proliferation of an organism. The invention also includes methods for identifying the target on which a compound which inhibits the proliferation of an organism acts and methods for identifying the extent to which a strain is present in a culture or collection of strains. Thus, a culture is obtained comprising a plurality of strains wherein each strain overexpresses a different gene product which is essential for proliferation. The culture is contacted with a sufficient concentration of an agent to inhibit the proliferation of strains which do not overexpress the gene product on which the agent acts, such that strains which overexpress the gene product on which the agent acts proliferate more rapidly than strains which do not overexpress said gene product on which the agent acts. The gene product which is overexpressed in a strain which proliferates more rapidly in the culture is then identified. Expression levels of gene transcripts are determined using hybridization and/or amplification methods standard to the art. Genes required for cellular proliferation of microbial organisms are identified by antisense RNA technology. Nucleotide sequences are provided for nucleic acid fragments whose expression results in detrimental effects on proliferation of *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhimurium*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, or *Enterococcus faecalis*. [This abstract record is one of three records for this document necessitated by the large number of index entries required to fully index the document and publication system constraints.]

IT 775416-81-4

RL: BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)

(amino acid sequence; target gene product identification for microbial cell proliferation-inhibiting compds.)

ACCESSION NUMBER: 2002:575103 CAPLUS Full-text
 DOCUMENT NUMBER: 137:168250
 TITLE: Hyperimmune serum-reactive antigens derived from
 expression libraries for treating or preventing
 pathogen infection, cancer, allergy, and autoimmune
 disease
 INVENTOR(S): Meinke, Andreas; Nagy, Eszter; Von Ahsen, Uwe; Klade,
 Christoph; Henics, Tamas; Zauner, Wolfgang; Minh, Duc
 Bui; Vytvytska, Oresta; Etz, Hildegard; Dryla,
 Agnieszka; Weichhart, Thomas; Hafner, Martin;
 Tempelmaier, Brigitte
 PATENT ASSIGNEE(S): Cistem Biotechnologies Gmbh, Austria; Intercell AG
 SOURCE: PCT Int. Appl., 252 pp.
 CODEN: PIXXD2
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 1
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2002059148	A2	20020801	WO 2002-EP546	20020121
WO 2002059148	C2	20021031		
W:				
AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN,				
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,				
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,				
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH,				
PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ,				
UA, UG, US, UZ, VN, YU, ZA, ZM, ZW				
RW:				
GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AM, AZ, BY,				
KG, KZ, MD, RU, TJ, TM, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB,				
GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA,				
GN, GQ, GW, ML, MR, NE, SN, TD, TG				
AT 200100130	A5	20021215	AT 2001-130	20010126
AT 410798	B	20030725		
CA 2436057	AA	20020801	CA 2002-2436057	20020121
EP 1355930	A2	20031029	EP 2002-716669	20020121
EP 1355930	B1	20051109		
R:				
AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT,				
IE, SI, LT, LV, FI, RO, MK, CY, AL, TR				
BR 2002007067	A	20040615	BR 2002-7067	20020121
JP 2004531476	T2	20041014	JP 2002-559450	20020121
NO 2003003364	A	20030924	NO 2003-3364	20030725
ZA 2003005764	A	20040726	ZA 2003-5764	20030725
US 2005037444	A1	20050217	US 2004-470048	20040206
PRIORITY APPLN. INFO.:			AT 2001-130	A 20010126
			WO 2002-EP546	W 20020121

ED Entered STN: 02 Aug 2002

AB Described is a method for identification, isolation and production of
 hyperimmune serum-reactive antigens from a specific pathogen, a tumor, an
 allergen or a tissue or host prone to autoimmunity that are suited for use as
 vaccines for treating related diseases in animals or humans. The method is
 characterized by providing an antibody preparation from a plasma pool of said
 given type of animal or from a human plasma pool or individual sera with
 antibodies against said specific pathogen, tumor, allergen or tissue or host
 prone to auto-immunity; providing at least one expression library of said
 specific pathogen, tumor, allergen or tissue or host prone to auto-immunity;
 screening said at least one expression library with said antibody preparation;

identifying antigens which bind in said screening to antibodies in said antibody preparation; screening the identified antigens with individual antibody preps. from individual sera from individuals with antibodies against said specific pathogen, tumor, allergen or tissue or host prone to auto-immunity; identifying the hyperimmune serum-reactive antigen portion of said identified antigens and which hyperimmune serum-reactive antigens bind to a relevant portion of said individual antibody preps. from said individual sera; and optionally isolating said hyperimmune serum-reactive antigens and producing said hyperimmune serum-reactive antigens by chemical or recombinant methods.

IT 445314-07-8P

RL: BPN (Biosynthetic preparation); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PREP (Preparation)
(amino acid sequence; hyperimmune serum-reactive antigens derived from expression libraries for treating or preventing pathogen infection, cancer, allergy, and autoimmune disease)

L6 ANSWER 6 OF 31 CAPLUS COPYRIGHT 2005 ACS on STN DUPLICATE 6

ACCESSION NUMBER: 2002:327941 CAPLUS Full-text

DOCUMENT NUMBER: 136:351426

TITLE: Nucleic acid and amino acid sequences relating to Staphylococcus epidermidis for diagnostics and therapeutics

INVENTOR(S): Doucette-Stamm, Lynn A.; Bush, David

PATENT ASSIGNEE(S): Genome Therapeutics Corporation, USA

SOURCE: U.S., 267 pp.
CODEN: USXXAM

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
-----	---	-----	-----	-----
US 6380370	B1	20020430	US 1998-134001	19980813
US 2004147734	A1	20040729	US 2003-724972	20031201
PRIORITY APPLN. INFO.:			US 1997-55779P	P 19970814
			US 1997-64964P	P 19971108
			US 1998-134001	A2 19980813
			US 1999-450969	A3 19991129

ED Entered STN: 02 May 2002

AB The invention provides isolated polypeptide and nucleic acid sequences derived from Staphylococcus epidermidis that are useful in diagnosis and therapy of pathol. conditions; antibodies against the polypeptides; and methods for the production of the polypeptides. Thus, the sequences of 2837 protein-coding contigs from the genome of S. epidermidis strain 19804 are provided. The invention also provides methods for the detection, prevention and treatment of pathol. conditions resulting from bacterial infection.

IT 421018-05-5P

RL: BPN (Biosynthetic preparation); BSU (Biological study, unclassified); DGN (Diagnostic use); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)

(amino acid sequence; nucleic acid and amino acid sequences relating to Staphylococcus epidermidis for diagnostics and therapeutics)

REFERENCE COUNT: 1 THERE ARE 1 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L6 ANSWER 7 OF 31 CAPLUS COPYRIGHT 2005 ACS on STN DUPLICATE 7

ACCESSION NUMBER: 2001:713538 CAPLUS Full-text

DOCUMENT NUMBER: 135:283990

TITLE: Identification of essential genes in prokaryotes and use of their antisense constructs in antibiotic screening

INVENTOR(S): Haselbeck, Robert; Ohlsen, Kari L.; Zyskind, Judith W.; Wall, Daniel; Trawick, John D.; Carr, Grant J.; Yamamoto, Robert T.; Xu, H. Howard

PATENT ASSIGNEE(S): Elitra Pharmaceuticals, Inc., USA

SOURCE: PCT Int. Appl., 511 pp.
CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001070955 A2		20010927	WO 2001-US9180	20010321
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG, TR			
PRIORITY APPLN. INFO.:			US 2000-PV191078	20000321
			US 2000-PV206848	20000523
			US 2000-PV207727	20000526
			US 2000-PV242578	20001023
			US 2000-PV253625	20001127
			US 2000-PV257931	20001222
			US 2001-PV269308	20010216

ED Entered STN: 28 Sep 2001

AB Genes required for proliferation of *Staphylococcus aureus*, *Salmonella typhimurium*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, and *Enterococcus faecalis*. Libraries of genomic fragments were operably cloned into vectors comprising inducible promoters in the antisense orientation, and selected for those genes which fail to grow or grow at a substantially reduced rate when the promoter is induced. The sequences of antisense nucleic acids which inhibit the proliferation of prokaryotes are disclosed. Cell-based assays which employ the antisense nucleic acids to identify and develop antibiotics are also disclosed. The antisense nucleic acids can also be used to identify proteins required for proliferation, express these proteins or portions thereof, obtain antibodies capable of specifically binding to the expressed proteins, and to use those expressed proteins as a screen to isolate candidate mols. for rational drug discovery programs. The nucleic acids can also be used to screen for homologous nucleic acids that are required for proliferation in cells other than *Staphylococcus aureus*, *Salmonella typhimurium*, *Klebsiella pneumoniae*, and *Pseudomonas aeruginosa*. The nucleic acids of the present invention can also be used in various assay systems to screen for proliferation required genes in other organisms.

IT 364143-92-0

RL: ARU (Analytical role, unclassified); BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); ANST (Analytical study); BIOL (Biological study); USES (Uses)

(amino acid sequence; identification of essential genes in prokaryotes and use of their antisense constructs in antibiotic screening)

DOCUMENT NUMBER: 132:330627
 TITLE: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum and proteins of said chromosome useful in anti-malarial vaccines and diagnostic reagents
 INVENTOR(S): Hoffman, Stephen; Carucci, Daniel; Gardner, Malcolm; Venter, J. Craig
 PATENT ASSIGNEE(S): USA
 SOURCE: PCT Int. Appl., 577 pp.
 CODEN: PIXXD2
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 1
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2000025728	A2	20000511	WO 1999-US26796	19991105
WO 2000025728	A3	20010222		
W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
AU 2000018182	A5	20000522	AU 2000-18182	19991105
PRIORITY APPLN. INFO.:			US 1998-107131P	P 19981105
			WO 1999-US26796	W 19991105

ED Entered STN: 15 May 2000

AB Chromosome 2 of Plasmodium falciparum was sequenced and shown to contain 945,000 base pairs and encode 209 predicted genes. Compared to the Saccharomyces cerevisiae genome, chromosome 2 has a lower gene d., introns are more frequent, and proteins are markedly enriched in non-globular domains. A new family of surface proteins, rifins, was identified. Rifins are believed to play a role in antigenic variation. The genome sequence provides a foundation for development of methods to control malaria, a disease that kills millions of people annually.

IT 257896-67-6

RL: ANT (Analyte); BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); ANST (Analytical study); BIOL (Biological study); OCCU (Occurrence); USES (Uses)
 (amino acid sequence; chromosome 2 sequence of the human malaria parasite Plasmodium falciparum and proteins of said chromosome useful in anti-malarial vaccines and diagnostic reagents)

L6 ANSWER 9 OF 31 CAPLUS COPYRIGHT 2005 ACS on STN

ACCESSION NUMBER: 2005:673319 CAPLUS Full-text

DOCUMENT NUMBER: 143:147803

TITLE: Spider dragline silk protein fusion with fibroin H-chain peptide produced by transposon-mediated transformation and production in silkworm

INVENTOR(S): Hiramatsu, Shingo; Moriyama, Hiromitsu; Asaoka, Ryota; Morita, Ken; Tanaka, Takashi; Yamada, Katsushige; Obrien, John Philip; Fahnestock, Stephen R.

PATENT ASSIGNEE(S): Toray Industries, Inc., Japan; E.I. Dupont de Nemours and Company

SOURCE: PCT Int. Appl., 48 pp.

DOCUMENT TYPE: Patent
LANGUAGE: Japanese
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2005068495	A1	20050728	WO 2005-JP619	20050112
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW			
RW:	BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG			

PRIORITY APPLN. INFO.: JP 2004-5489 A 20040113

ED Entered STN: 29 Jul 2005

AB The invention relates to novel silk containing spider dragline silk protein. Using transgenic silkworms transformed with a gene encoding a spider dragline silk protein having desired properties (a high strength, a high elongation, etc.), a hybrid silk of spider dragline silk with silk thread having the desired properties is produced. Transformation can be achieved without damaging silkworm fibroin H-chain gene using transposons,. The spider dragline silk protein is produced as fusion with fibroin H-chain peptide, and forms a disulfide linkage with fibroin L-chain via C-terminal cysteine. Synthetic spider dragline silk proteins containing repeats of DP-1B.33 (dragline protein 1 analog) were produced and fused with silkworm fibroin H-chain using piggyBac transposon.

IT 859621-51-5

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)

(amino acid sequence; spider dragline silk protein fusion with fibroin H-chain peptide produced by transposon-mediated transformation and production in silkworm)

REFERENCE COUNT: 10 THERE ARE 10 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L6 ANSWER 10 OF 31 CAPLUS COPYRIGHT 2005 ACS on STN

ACCESSION NUMBER: 2004:260861 CAPLUS Full-text

DOCUMENT NUMBER: 140:265688

TITLE: Soybean nucleic acids and encoded proteins associated with transcription in plants and their uses for plant improvement

INVENTOR(S): La Rosa, Thomas J.; Zhou, Yihua; Kovalic, David K.; Cao, Yongwei

PATENT ASSIGNEE(S): USA

SOURCE: U.S. Pat. Appl. Publ., 15 pp., Cont.-in-part of U.S. Ser. No. 985,678, abandoned.

CODEN: USXXCO

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 76

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
------------	------	------	-----------------	------

-----	-----	-----	-----	-----
US 2004031072	A1	20040212	US 2003-424599	20030428
US 2004031072	A1	20040212	US 2003-424599	20030428
PRIORITY APPLN. INFO.:			US 1999-304517	B1 19990506
			US 2001-985678	B2 20011105
			US 2003-424599	A 20030428

ED Entered STN: 31 Mar 2004

AB This invention provides 142,842 polynucleotide sequences isolated from a cDNA library generated from Glycine maximum. The open reading frame in each polynucleotide sequence is identified by a combination of predictive and homol.-based methods. Functions of polypeptides encoded by the polynucleotides sequences are determined using a hierarchical classification tool, termed FunCAT, for Functional Categories Annotation Tool. Sequences useful for producing transgenic plants having improved biol. properties are identified from their FunCAT annotations. [This abstract record is one of 72 records for this document necessitated by the large number of index entries required to fully index the document and publication system constraints.].

IT 672991-01-4

RL: BSU (Biological study, unclassified); BUU (Biological use, unclassified); PRP (Properties); BIOL (Biological study); USES (Uses) (amino acid sequence; soybean nucleic acids and encoded proteins associated with transcription in plants and their uses for plant improvement)

L6 ANSWER 11 OF 31 CAPLUS COPYRIGHT 2005 ACS on STN

ACCESSION NUMBER: 2004:1035641 CAPLUS Full-text

DOCUMENT NUMBER: 142:33017

TITLE: Cell growth-promoting peptides from silk proteins

INVENTOR(S): Tsubouchi, Kozo; Yamada, Hiroo

PATENT ASSIGNEE(S): National Institute of Agrobiological Resources NIAR, Japan

SOURCE: Jpn. Kokai Tokkyo Koho, 27 pp.

CODEN: JKXXAF

DOCUMENT TYPE: Patent

LANGUAGE: Japanese

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
-----	----	-----	-----	-----
JP 2004339189	A2	20041202	JP 2003-406608	20031204
US 2005143296	A1	20050630	US 2004-789494	20040227
CN 1535723	A	20041013	CN 2004-10035241	20040301
PRIORITY APPLN. INFO.:			JP 2003-55048	A 20030228

ED Entered STN: 03 Dec 2004

AB Disclosed are cell growth-promoting peptides which comprise 4-40 amino acids from noncryst. peptide chains of the silk proteins. The peptides are obtained by hydrolyzing silk worm proteins or Antheraea cocoon fibroins and separating them by mol. weight fraction. The peptides are effective as cell growth promoters, cell adhesives, wound healing promoters, and cell culture matrixes. Also claimed is a cosmetic containing the peptides.

IT 714954-21-9P

RL: COS (Cosmetic use); NPO (Natural product occurrence); PAC (Pharmacological activity); PNU (Preparation, unclassified); BIOL (Biological study); OCCU (Occurrence); PREP (Preparation); USES (Uses) (cell growth-promoting peptides from silk proteins)

IT 803823-75-8 803823-77-0

RL: PRP (Properties)

(unclaimed sequence; cell growth-promoting peptides from silk proteins)

L6 ANSWER 12 OF 31 CAPLUS COPYRIGHT 2005 ACS on STN

ACCESSION NUMBER: 2003:838448 CAPLUS Full-text

DOCUMENT NUMBER: 141:82207

TITLE: Identification of fibroin-derived peptides enhancing the proliferation of cultured human skin fibroblasts

AUTHOR(S): Yamada, Hiromi; Igarashi, Yumiko; Takasu, Yoko; Saito, Hitoshi; Tsubouchi, Kozo

CORPORATE SOURCE: Entomological Science, National Institute of Agrobiological Sciences, Tsukuba, Ibaraki, 305-8634, Japan

SOURCE: Biomaterials (2003), Volume Date 2004, 25(3), 467-472

CODEN: BIMADU; ISSN: 0142-9612

PUBLISHER: Elsevier Science Ltd.

DOCUMENT TYPE: Journal

LANGUAGE: English

ED Entered STN: 27 Oct 2003

AB The authors previously reported that the fibroin of the silkworm *Bombyx mori* enhanced the proliferation of cultured human skin fibroblasts. In this work, the fibroin was digested by chymotrypsin, and the resulting peptide fragments were fractionated and assayed for their biol. activity. Two peptides that promoted fibroblast growth were isolated and identified to be VITDSDGNE and NINDFDED. Both sequences are found in the N-terminal region of the fibroin polypeptide and are thought to be the active principle of fibroblast growth-promoting activity.

IT 714954-21-9

RL: PAC (Pharmacological activity); THU (Therapeutic use); BIOL

(Biological study); USES (Uses)

(fibroin-derived peptides enhancing proliferation of cultured human skin fibroblasts)

REFERENCE COUNT: 12 THERE ARE 12 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L6 ANSWER 13 OF 31 CAPLUS COPYRIGHT 2005 ACS on STN

ACCESSION NUMBER: 2003:593459 CAPLUS Full-text

DOCUMENT NUMBER: 139:287022

TITLE: The 62-kb upstream region of *Bombyx mori* fibroin heavy chain gene is clustered of repetitive elements and candidate matrix association regions

AUTHOR(S): Zhou, Cong-Zhao; Confalonieri, Fabrice; Esnault, Catherine; Zivanovic, Yvan; Jacquet, Michel; Janin, Joel; Perasso, Roland; Li, Zhen-Gang; Duguet, Michel

CORPORATE SOURCE: Institut de Genetique et Microbiologie, Universite Paris-Sud et CNRS, Orsay, 91405, Fr.

SOURCE: Gene (2003), 312, 189-195

CODEN: GENED6; ISSN: 0378-1119

PUBLISHER: Elsevier Science B.V.

DOCUMENT TYPE: Journal

LANGUAGE: English

ED Entered STN: 04 Aug 2003

AB We sequenced an 80 kb DNA region containing the complete sequence of the silkworm *Bombyx mori* fibroin gene and its flanking, especially the upstream, regions (.apprx.62 kb). About 30% of the 62 kb upstream region is composed of repetitive elements including short interspersed elements Bm1, long interspersed elements L1Bm and mariner-like elements Bmmar1 which are widespread over the silkworm genome. This 62 kb region is also enriched of commonly considered matrix association region (MAR) motifs. A total of 25 individual MAR recognition signatures (MRSs) were identified, with 24 at the upstream and one at the downstream region. Combining two newly developed MAR prediction programs (MAR-finder and Chrclass), ten candidate MARs were predicted, with five containing MRS and seven related to the repetitive

elements. The wide distribution of nested repetitive elements, candidate MARS, DNase I hypersensitive sites and other potential regulatory factors recognition sites indicates this region is probably a unique huge cis-acting element contributing to the regulation of the spatial and temporal specificity and efficiency of fibroin gene expression.

IT 303229-60-9, Fibroin heavy chain (silkworm strain p50)

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL
(Biological study)

(amino acid sequence; 62-kb upstream region of Bombyx mori fibroin heavy chain gene has clustered repetitive elements and candidate matrix association regions)

REFERENCE COUNT: 33 THERE ARE 33 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L6 ANSWER 14 OF 31 CAPLUS COPYRIGHT 2005 ACS on STN

ACCESSION NUMBER: 2002:752115. CAPLUS Full-text

DOCUMENT NUMBER: 137:289734

TITLE: Sequence of Plasmodium falciparum chromosomes 2, 10, 11 and 14

AUTHOR(S): Gardner, Malcolm J.; Shallom, Shamira J.; Carlton, Jane M.; Salzberg, Steven L.; Nene, Vishvanath; Shoaibi, Azadeh; Ciecko, Anne; Lynn, Jeffery; Rizzo, Michael; Weaver, Bruce; Jarrahi, Behnam; Brenner, Michael; Parvizi, Babak; Tallon, Luke; Moazzez, Azita; Granger, David; Fujii, Claire; Hansen, Cheryl; Pederson, James; Feldblyum, Tamara; Peterson, Jeremy; Suh, Bernard; Angiuoli, Sam; Perte, Mihaela; Allen, Jonathan; Selengut, Jeremy; White, Owen; Cummings, Leda M.; Smith, Hamilton O.; Adams, Mark D.; Venter, J. Craig; Carucci, Daniel J.; Hoffman, Stephen L.; Fraser, Claire M.

CORPORATE SOURCE: The Institute for Genomic Research, Rockville, MD, 20850, USA

SOURCE: Nature (London, United Kingdom) (2002), 419(6906), 531-534

CODEN: NATUAS; ISSN: 0028-0836

PUBLISHER: Nature Publishing Group

DOCUMENT TYPE: Journal

LANGUAGE: English

ED Entered STN: 04 Oct 2002

AB The mosquito-borne malaria parasite Plasmodium falciparum kills an estimated 0.7-2.7 million people every year, primarily children in sub-Saharan Africa. Without effective interventions, a variety of factors-including the spread of parasites resistant to antimalarial drugs and the increasing insecticide resistance of mosquitoes-may cause the number of malaria cases to double over the next two decades. To stimulate basic research and facilitate the development of new drugs and vaccines, the genome of Plasmodium falciparum clone 3D7 has been sequenced using a chromosome-by-chromosome shotgun strategy. This report describes nucleotide sequences of chromosomes 10, 11 and 14, and a re-anal. of the chromosome 2 sequence. These chromosomes represent about 35% of the 23-megabase P. falciparum genome. The sequences are deposited in GenBank/EMBL/DBJ under accession nos. AE001362.2 (chromosome 2), AE014185 (chromosome 10), AE014186 (chromosome 11), and AE014187 (chromosome 14).

IT 465598-80-5 465605-62-3

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL
(Biological study)

(amino acid sequence; complete sequence of Plasmodium falciparum chromosomes 2, 10, 11 and 14)

REFERENCE COUNT: 30 THERE ARE 30 CITED REFERENCES AVAILABLE FOR THIS

RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L6 ANSWER 15 OF 31 CAPLUS COPYRIGHT 2005 ACS on STN

ACCESSION NUMBER: 2002:752116 CAPLUS Full-text

DOCUMENT NUMBER: 137:289735

TITLE: Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13

AUTHOR(S): Hall, N.; Pain, A.; Berriman, M.; Churcher, C.; Harris, B.; Harris, D.; Mungall, K.; Bowman, S.; Atkin, R.; Baker, S.; Barron, A.; Brooks, K.; Buckee, C. O.; Burrows, C.; Cherevach, I.; Chillingworth, C.; Chillingworth, T.; Christodoulou, Z.; Clark, L.; Clark, R.; Corton, C.; Cronin, A.; Davies, R.; Davis, P.; Dear, P.; Dearden, F.; Doggett, J.; Feltwell, T.; Goble, A.; Goodhead, I.; Gwilliam, R.; Hamlin, N.; Hance, Z.; Harper, D.; Hauser, H.; Hornsby, T.; Holroyd, S.; Horrocks, P.; Humphray, S.; Jagels, K.; James, K. D.; Johnson, D.; Kerhornou, A.; Knights, A.; Konfortov, B.; Kyes, S.; Larke, N.; Lawson, D.; Lennard, N.; Line, A.; Maddison, M.; McLean, J.; Mooney, P.; Moule, S.; Murphy, L.; Oliver, K.; Ormond, D.; Price, C.; Quail, M. A.; Rabinowitsch, E.; Rajandream, M.-A.; Rutter, S.; Rutherford, K. M.; Sanders, M.; Simmonds, M.; Seeger, K.; Sharp, S.; Smith, R.; Squares, R.; Squares, S.; Stevens, K.; Taylor, K.; Tivey, A.; Unwin, L.; Whitehead, S.; Woodward, J.; Sulston, J. E.; Craig, A.; Newbold, C.; Barrell, B. G.

CORPORATE SOURCE: The Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, UK

SOURCE: Nature (London, United Kingdom) (2002), 419(6906), 527-531

CODEN: NATUAS; ISSN: 0028-0836

PUBLISHER: Nature Publishing Group

DOCUMENT TYPE: Journal

LANGUAGE: English

ED Entered STN: 04 Oct 2002

AB Since the sequencing of the first two chromosomes of the malaria parasite, Plasmodium falciparum, there has been a concerted effort to sequence and assemble the entire genome of this organism. This report provides the sequence of chromosomes 1, 3-9 and 13 of P. falciparum clone 3D7; these chromosomes account for .apprx.55% of the total genome. The methods used to map, sequence and annotate these chromosomes is described. By comparing these assemblies with the optical map, the completeness of the resulting sequence is indicated. During annotation, Gene Ontol. terms were assigned to the predicted gene products, and clustering of some malaria-specific terms to specific chromosomes was observed. A highly conserved sequence element was found in the intergenic region of internal var genes that is not associated with their telomeric counterparts.

IT 467525-63-9

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)

(amino acid sequence; sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13)

REFERENCE COUNT: 33 THERE ARE 33 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L6 ANSWER 16 OF 31 CAPLUS COPYRIGHT 2005 ACS on STN

ACCESSION NUMBER: 2002:235070 CAPLUS Full-text

DOCUMENT NUMBER: 137:289612

TITLE: Cloning and structure analysis on 5' flanking sequence
of fibroin gene of chinese oak silkworm
AUTHOR(S): Li, Wenli; Jin, Liji; Fan, Qi; An, Lijia
CORPORATE SOURCE: Department of Biotechnology, Dalian University of
Technology, Dalian, 116023, Peop. Rep. China
SOURCE: Zhongguo Nongye Kexue (Beijing, China) (2002), 35(2),
218-221
CODEN: CKNYAR; ISSN: 0578-1752
PUBLISHER: Zhongguo Nongye Kexue Bianjibu
DOCUMENT TYPE: Journal
LANGUAGE: Chinese

ED Entered STN: 28 Mar 2002

AB The 5' flanking fragment of fibroin gene of Chinese Oak Silkworm (*Antheraea pernyi*) was amplified through PCR. It consists of CAAT box, TATA box(Hogness box), prim transcript, start code ATG, part of structural gene and first intron. Compared with Japanese Oak silkworm (*Antheraea yamamai*), three high homol. regions have been observed in the sequence of the 5' flanking (nt. 86.apprx.479bp, number 769.apprx.1167bp and nt. 1189.apprx.1303bp), with the similarity of 91.6%, 95% and 95%, resp. As for CAAT box, TATA box and prim transcript, the homol. between Chinese Oak Silkworm and Japanese Oak silkworm is higher than that between Chinese Oak silkworm and *Bombyx mori*. The TATA box locate at the upstream -25bp and CAAT box at - 70bp from the prim transcript, which is similar as the character of eukarytor promoter.

IT 469866-34-0

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL
(Biological study)
(amino acid sequence; cloning and structure anal. on 5' flanking
sequence of fibroin gene of chinese oak silkworm)

L6 ANSWER 17 OF 31 CAPLUS COPYRIGHT 2005 ACS on STN

ACCESSION NUMBER: 2001:360182 CAPLUS Full-text
DOCUMENT NUMBER: 135:1253
TITLE: Staphylococcus epidermidis nucleic acids and proteins
as diagnostic or therapeutic agents
INVENTOR(S): Kimmerly, William John
PATENT ASSIGNEE(S): Glaxo Group Limited, UK
SOURCE: PCT Int. Appl., 2189 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001034809	A2	20010517	WO 2000-US30782	20001109
WO 2001034809	A3	20011122		
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW				
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
US 6703492	B1	20040309	US 2000-710279	20001109
US 2005255478	A1	20051117	US 2004-793626	20040304
PRIORITY APPLN. INFO.:			US 1999-164258P	P 19991109
			US 2000-710279	A3 20001109

ED Entered STN: 18 May 2001

AB Staphylococcus epidermidis polypeptides and DNA (RNA) encoding such polypeptides and a procedure for producing such polypeptides by recombinant techniques is disclosed. The sequences of 1667 genes and their encodes proteins, as well as 1120 noncoding nucleic acid fragments, are provided. Also disclosed are methods for utilizing such polypeptides and DNA (RNA) for the treatment of infection, particularly infections arising from S. epidermidis. Antagonists against the function of such polypeptides and their use as therapeutics to treat infection are also disclosed. Selected nucleic acids and/or polypeptides of the present invention can be advantageously utilized as targets in screening assays for antibiotics, as diagnostics of infections, and as means to identify S. epidermidis in any given sample and distinguish it from other bacteria.

IT 341040-34-4

RL: ANT (Analyte); PRP (Properties); THU (Therapeutic use); ANST (Analytical study); BIOL (Biological study); USES (Uses)
(amino acid sequence; Staphylococcus epidermidis nucleic acids and proteins as diagnostic or therapeutic agents)

L6 ANSWER 18 OF 31 CAPLUS COPYRIGHT 2005 ACS on STN

ACCESSION NUMBER: 2001:669212 CAPLUS Full-text

DOCUMENT NUMBER: 136:242651

TITLE: Cloning of the fibroin gene from the oak silkworm, Antheraea yamamai and its complete sequence

AUTHOR(S): Hwang, Jae-Sam; Lee, Jin-Sung; Goo, Tae-Won; Yun, Eun-Young; Lee, Kwang-Sik; Kim, Yong-Sung; Jin, Byung-Rae; Lee, Sang-Mong; Kim, Keun-Young; Kang, Seok-Woo; Suh, Dong-Sang

CORPORATE SOURCE: Department of Sericulture and Entomology, National Institute of Agricultural Science and Technology, RDA, Suwon, 441-100, S. Korea

SOURCE: Biotechnology Letters (2001), 23(16), 1321-1326

CODEN: BILED3; ISSN: 0141-5492

PUBLISHER: Kluwer Academic Publishers

DOCUMENT TYPE: Journal

LANGUAGE: English

ED Entered STN: 13 Sep 2001

AB The nucleotide sequences containing an entire genomic region and 5' upstream region of Antheraea yamamai fibroin gene have been determined. The gene consists of an initial exon encoding 14 amino acids, an intron (150 bp), and a long second exon coding for 2641 amino acids. The fibroin coding sequence shows a specialized organization with a highly repetitive region flanked by non repetitive 5' and 3' ends. Northern blot analyses confirmed that fibroin gene is actively expressed in the posterior silk gland of the final instar larvae of Antheraea yamamai.

IT 404318-03-2, Fibroin (Antheraea yamamai)

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)

(amino acid sequence; sequence of the fibroin gene from the oak silkworm, Antheraea yamamai)

REFERENCE COUNT: 12 THERE ARE 12 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L6 ANSWER 19 OF 31 CAPLUS COPYRIGHT 2005 ACS on STN

ACCESSION NUMBER: 2000:472155 CAPLUS Full-text

DOCUMENT NUMBER: 133:330213

TITLE: Fine organization of Bombyx mori fibroin heavy chain gene

AUTHOR(S): Zhou, Cong-Zhao; Confalonieri, Fabrice; Medina, Nadine; Zivanovic, Yvan; Esnault, Catherine; Yang,

Tie; Jacquet, Michel; Janin, Joel; Duguet, Michel;
Perasso, Roland; Li, Zhen-Gang
CORPORATE SOURCE: Institut de Genetique et Microbiologie and Laboratoire
de Biologie Cellulaire 4, Universite Paris-Sud et
CNRS, Orsay, 91405, Fr.
SOURCE: Nucleic Acids Research (2000), 28(12), 2413-2419
CODEN: NARHAD; ISSN: 0305-1048
PUBLISHER: Oxford University Press
DOCUMENT TYPE: Journal
LANGUAGE: English

ED Entered STN: 13 Jul 2000

AB The complete sequence of the Bombyx mori fibroin gene has been determined by means of combining a shotgun sequencing strategy with phys. map-based sequencing procedures. It consists of two exons (67 and 15 750 bp, resp.) and one intron (971 bp). The fibroin coding sequence presents a spectacular organization, with a highly repetitive and G-rich (.apprx.45%) core flanked by non-repetitive 5' and 3' ends. This repetitive core is composed of alternate arrays of 12 repetitive and 11 amorphous domains. The sequences of the amorphous domains are evolutionarily conserved and the repetitive domains differ from each other in length by a variety of tandem repeats of subdomains of .apprx.208 bp which are reminiscent of the repetitive nucleosome organization. A typical composition of a subdomain is a cluster of repetitive units, Ua, followed by a cluster of units, Ub, (with a Ua:Ub ratio of 2:1) flanked by conserved boundary elements at the 3' end. Moreover some repeats are also perfectly conserved at the peptide level indicating that the evolutionary pressure is not identical along the sequence. A tentative model for the constitution and evolution of this unusual gene is discussed.

IT 303229-60-9

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL
(Biological study)

(amino acid sequence; fine organization of Bombyx mori fibroin heavy chain gene)

REFERENCE COUNT: 29 THERE ARE 29 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L6 ANSWER 20 OF 31 CAPLUS COPYRIGHT 2005 ACS on STN

ACCESSION NUMBER: 2000:831431 CAPLUS Full-text

DOCUMENT NUMBER: 134:362047

TITLE: Dynamic rearrangement within the Antheraea pernyi silk fibroin gene is associated with four types of repetitive units

AUTHOR(S): Sezutsu, Hideki; Yukuhiro, Kenji

CORPORATE SOURCE: Department of Insect Genetic Breeding, National Institute of Sericultural and Entomological Science, Tsukuba, 305-8634, Japan

SOURCE: Journal of Molecular Evolution (2000), 51(4), 329-338
CODEN: JMEVAU; ISSN: 0022-2844

PUBLISHER: Springer-Verlag New York Inc.

DOCUMENT TYPE: Journal

LANGUAGE: English

ED Entered STN: 29 Nov 2000

AB We characterized a full-length gene encoding wild silkworm Antheraea pernyi fibroin (Ap-fibroin) to clarify the conformation of repetitive sequences. The gene consisted of a first exon encoding 14 amino acid residues, a short intron (120 bp), and a long second exon encoding 2,625 amino acid residues. Three amino acids, alanine, glycine, and serine, amounted to 81% of the Ap-fibroin sequence. The Ap-fibroin, except for 155 residues of the amino terminus, was composed of 80 tandemly arranged polyalanine-containing units (motifs). A motif was a doublet of a polyalanine block (PAB) and a nonpolyalanine block (NPAB). Seventy-eight of the 80 motifs were classified into four types based

on differences in the NPAB sequences. Although resp. motifs were significantly conserved, many rearrangements were observed within the second exon, i.e., the triplication of a 558-bp-long sequence and other duplication events of shorter sequences. Chi-like sequences, GCTGGAG, might contribute to the rearrangement within the gene as described in human minisatellite loci, because they were found at specific sites of NPAB-encoding sequences in three of four types of motifs. The present results support the idea that the Ap-fibroin gene is unstable like minisatellite sequences and that the evolution of this gene is strongly associated with its instability.

IT 336885-96-2, Fibroin (Antheraea pernyi clone AP2)

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL
(Biological study)

(amino acid sequence; dynamic rearrangement within the Antheraea pernyi silk fibroin gene is associated with four types of repetitive units)

REFERENCE COUNT: 24 THERE ARE 24 CITED REFERENCES AVAILABLE FOR THIS
RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L6 ANSWER 21 OF 31 CAPLUS COPYRIGHT 2005 ACS on STN

ACCESSION NUMBER: 1999:590975 CAPLUS Full-text

DOCUMENT NUMBER: 132:147372

TITLE: The cytoadherence linked asexual gene family of
Plasmodium falciparum: are there roles other than
cytoadherence?

AUTHOR(S): Holt, Deborah C.; Gardiner, Donald L.; Thomas,
Elizabeth A.; Mayo, Mark; Bourke, Peter F.;
Sutherland, Colin J.; Carter, Rod; Myers, Garry; Kemp,
David J.; Trenholme, Katharine R.

CORPORATE SOURCE: The Menzies School of Health Research, Casuarina,
0811, Australia

SOURCE: International Journal for Parasitology (1999), 29(6),
939-944

CODEN: IJPYBT; ISSN: 0020-7519

PUBLISHER: Elsevier Science Ltd.

DOCUMENT TYPE: Journal

LANGUAGE: English

ED Entered STN: 21 Sep 1999

AB The binding of erythrocytes infected with P. falciparum to the endothelium lining the small blood vessels of the brain and other organs can mediate severe pathol. A region at the right end of chromosome 9 has been implicated in the binding of parasitized erythrocytes to the endothelial receptor CD36. A gene expressed in asexual erythrocytic stage parasites has been identified in this region and termed the cytoadherence linked asexual gene (clag). Antisense RNA production and targeted gene disruption of clag resulted in greatly reduced binding to CD36. Hybridization to 3D7 chromosomes showed clag to be a part of a gene family of at least nine members. All members analyzed so far have a conserved gene structure of at least nine exons, as well as putative transmembrane domains. The possible functions of the gene family are discussed.

IT 257896-67-6

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL
(Biological study)

(amino acid sequence; cytoadherence linked asexual gene family of
Plasmodium falciparum)

REFERENCE COUNT: 25 THERE ARE 25 CITED REFERENCES AVAILABLE FOR THIS
RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L6 ANSWER 22 OF 31 CAPLUS COPYRIGHT 2005 ACS on STN

ACCESSION NUMBER: 2001:364016 CAPLUS Full-text

DOCUMENT NUMBER: 135:1093

TITLE: The malaria genome sequencing project: Complete

AUTHOR(S): sequence of Plasmodium falciparum chromosome 2
Gardner, M. J.; Tettelin, H.; Carucci, D. J.;
Cummings, L. M.; Smith, H. O.; Fraser, C. M.; Venter,
J. C.; Hoffman, S. L.
CORPORATE SOURCE: The Institute for Genomic Research, Rockville, MD,
20850, USA
SOURCE: Parassitologia (Roma, Italy) (1999), 41(1-3), 69-75
CODEN: PSSGAR; ISSN: 0048-2951
PUBLISHER: Lambardo Editore
DOCUMENT TYPE: Journal
LANGUAGE: English

ED Entered STN: 20 May 2001

AB An international consortium has been formed to sequence the entire genome of the human malaria parasite Plasmodium falciparum. Chromosome 2 of clone 3D7 was sequenced using a shotgun sequencing strategy. Chromosome 2 is 947 kb in length, has a base composition of 80.2% A+T, and contains 210 predicted genes. In comparison to the Saccharomyces cerevisiae genome, chromosome 2 has a lower gene d., a greater proportion of genes containing introns, and nearly twice as many proteins containing predicted non-globular domains. A group of putative surface proteins was identified, rifins, which are encoded by a gene family comprising up to 7% of the protein-encoding genes in the genome. The rifins exhibit considerable sequence diversity and may play an important role in antigenic variation. Sixteen genes encoded on chromosome 2 showed signs of a plastid or mitochondrial origin, including several genes involved in fatty acid biosynthesis. Completion of the chromosome 2 sequence demonstrated that the A+T-rich genome of P. falciparum can be sequenced by the shotgun approach. Within 2-3 yr, the sequence of almost all P. falciparum genes will have been determined, paving the way for genetic, biochem. and immunol. research aimed at developing new drugs and vaccines against malaria.

IT 257896-67-6

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL
(Biological study)
(amino acid sequence; complete sequence of Plasmodium falciparum
chromosome 2)

REFERENCE COUNT: 43 THERE ARE 43 CITED REFERENCES AVAILABLE FOR THIS
RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L6 ANSWER 23 OF 31 USPATFULL on STN

ACCESSION NUMBER: 2005:165871 USPATFULL Full-text
TITLE: Extraction and utilization of cell growth-promoting
peptides from silk protein
INVENTOR(S): Tsubouchi, Kozo, Tsukuba-gun, JAPAN
Yamada, Hiromi, Tsukuba-shi, JAPAN

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2005143296	A1	20050630
APPLICATION INFO.:	US 2004-789494	A1	20040227 (10)

	NUMBER	DATE
PRIORITY INFORMATION:	JP 2003-55048	20030228
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	APPLICATION	
LEGAL REPRESENTATIVE:	FLYNN THIEL BOUTELL & TANIS, P.C., 2026 RAMBLING ROAD, KALAMAZOO, MI, 49008-1699, US	
NUMBER OF CLAIMS:	15	
EXEMPLARY CLAIM:	1	
LINE COUNT:	1910	

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

AB Peptides are provided having an excellent safety, stability due to relatively low molecular weights thereof, and cell growth promotion, which are different from cell growth factors produced by abnormal cells such as tumor cells. Peptide compositions which are excellent in promoting cell growth containing partial peptides of one or more peptide chains selected from peptide chains forming noncrystalline portions constituting silk protein. The partial peptides have specific amino acid sequences formed of four to forty amino acid residues. This invention has succeeded in providing novel peptides excellent for cell growth by separating and fractionating peptides, having specific amino acid sequences of molecular weights not higher than 10,000, preferably ranging from 4,000 to 400, from the noncrystalline portions of silk protein as well as by synthesizing peptides similar to such peptides. These peptides may be used for biomaterials such as a cell adhesion agent, cell growth-promoting agent, wound healing promoting agent, skin care material like cosmetic material or the like, and cell culture substrate.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

IT 714954-21-9P

(cell growth-promoting peptides from silk proteins)

IT 803823-75-8 803823-77-0

(unclaimed sequence; cell growth-promoting peptides from silk proteins)

L6 ANSWER 24 OF 31 USPATFULL on STN

ACCESSION NUMBER: 2005:158196 USPATFULL Full-text

TITLE: Nucleic acid and amino acid sequences relating to streptococcus pneumoniae for diagnostics and therapeutics

INVENTOR(S): Doucette-Stamm, Lynn A., Framingham, MA, UNITED STATES
Bush, David, Somerville, MA, UNITED STATES

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2005136404	A1	20050623
APPLICATION INFO.:	US 2003-617320	A1	20030710 (10)
RELATED APPLN. INFO.:	Division of Ser. No. US 1998-107433, filed on 30 Jun 1998, PENDING		

	NUMBER	DATE
PRIORITY INFORMATION:	US 1997-51553P	19970702 (60)
	US 1998-85131P	19980512 (60)
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	APPLICATION	
LEGAL REPRESENTATIVE:	Robert L. Spadafora, Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA, 02453, US	
NUMBER OF CLAIMS:	28	
EXEMPLARY CLAIM:	1	
LINE COUNT:	12957	

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

AB The invention provides isolated polypeptide and nucleic acid sequences derived from Streptococcus pneumonia that are useful in diagnosis and therapy of pathological conditions; antibodies against the polypeptides; and methods for the production of the polypeptides. The invention also provides methods for the detection, prevention and treatment of pathological conditions resulting from bacterial infection.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

IT 660056-86-0

(amino acid sequence; nucleic acid and amino acid sequences relating to Streptococcus pneumoniae for diagnostics and therapeutics)

L6 ANSWER 25 OF 31 USPATFULL on STN

ACCESSION NUMBER: 2005:43716 USPATFULL Full-text

TITLE: Method for identification isolation and production of antigens to a specific pathogen

INVENTOR(S): Meinke, Andreas, Pressbaum, AUSTRIA
Nagy, Eszter, Vienna, AUSTRIA
von Ahsen, Uwe, Vienna, AUSTRIA
Klade, Christoph, Neustadt, AUSTRIA
Henics, Tamas, Vienna, AUSTRIA
Zauner, Wolfgang, Vienna, AUSTRIA
Minh, Duc Bui, Vienna, AUSTRIA
Vytvytska, Oresta, Vienna, AUSTRIA
Etz, Hildegard, Vienna, AUSTRIA
Dryla, Agnieszka, Vienna, AUSTRIA
Weichhart, Thomas, Boheimkirchen, AUSTRIA
Hafner, Martin, Vienna, AUSTRIA
Tempelmaier, Brigitte, Vienna, AUSTRIA
Fraser, Claire M., Potomac, MD, UNITED STATES
Gill, Steven, Frederick, MD, UNITED STATES

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2005037444	A1	20050217
APPLICATION INFO.:	US 2004-470048	A1	20040206 (10)
	WO 2002-EP546		20020121

	NUMBER	DATE
PRIORITY INFORMATION:	AT 2001-130	20010126
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	APPLICATION	
LEGAL REPRESENTATIVE:	Mark B Wilson, Fulbright & Jaworski, Suite 2400, 600 Congress Avenue, Austin, TX, 78701	
NUMBER OF CLAIMS:	48	
EXEMPLARY CLAIM:	CLM-01-34	
NUMBER OF DRAWINGS:	11 Drawing Page(s)	
LINE COUNT:	4784	

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

AB Described is a method for identification, isolation and production of hyperimmune serum-reactive antigens from a specific pathogen, a tumor, an allergen or a tissue or host prone to autoimmunity, said antigens being suited for use in a vaccine for a given type of animal or for humans, which is characterized by the following steps:--providing an antibody preparation from a plasma pool of said given type of animal or from a human plasma pool or individual sera with antibodies against said specific pathogen, tumor, allergen or tissue or host prone to auto-immunity,--providing at least one expression library of said specific pathogen, tumor, allergen or tissue or host prone to auto-immunity,--screening said at least one expression library with said antibody preparation,--identifying antigens which bind in said screening to antibodies in said antibody preparation,--screening the identified antigens with individual antibody preparations from individual sera from individuals with antibodies against said specific pathogen, tumor, allergen or tissue or host prone to auto-immunity,--identifying the hyperimmune serum-reactive antigen portion of said identified antigens and which hyperimmune serum-reactive antigens bind to a relevant portion of said individual antibody preparations from said individual sera and--optionally

isolating said hyperimmune serum-reactive antigens and producing said hyperimmune serum-reactive antigens by chemical or recombinant methods.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

IT 445314-07-8P

(amino acid sequence; hyperimmune serum-reactive antigens derived from expression libraries for treating or preventing pathogen infection, cancer, allergy, and autoimmune disease)

L6 ANSWER 26 OF 31 USPATFULL on STN

ACCESSION NUMBER: 2004:190960 USPATFULL Full-text

TITLE: Nucleic acid and amino acid sequences relating to staphylococcus epidermidis for diagnostics and therapeutics

INVENTOR(S): Doucette-Stamm, Lynn, Framingham, MA, UNITED STATES
Bush, David, Somerville, MA, UNITED STATES

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2004147734	A1	20040729
APPLICATION INFO.:	US 2003-724972	A1	20031201 (10)
RELATED APPLN. INFO.:	Division of Ser. No. US 1999-450969, filed on 29 Nov 1999, PENDING Continuation-in-part of Ser. No. US 1998-134001, filed on 13 Aug 1998, GRANTED, Pat. No. US 6380370		

	NUMBER	DATE
PRIORITY INFORMATION:	US 1997-64964P	19971108 (60)
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	APPLICATION	
LEGAL REPRESENTATIVE:	OSCIENT THERAPEUTICS CORPORATION, 100 BEAVER STREET, WALTHAM, MA, 02453	
NUMBER OF CLAIMS:	31	
EXEMPLARY CLAIM:	1	
LINE COUNT:	3207	

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

AB The invention provides isolated polypeptide and nucleic acid sequences derived from Staphylococcus epidermidis that are useful in diagnosis and therapy of pathological conditions; antibodies against the polypeptides; and methods for the production of the polypeptides. The invention also provides methods for the detection, prevention and treatment of pathological conditions resulting from bacterial infection.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

IT 421018-05-5P

(amino acid sequence; nucleic acid and amino acid sequences relating to Staphylococcus epidermidis for diagnostics and therapeutics)

L6 ANSWER 27 OF 31 USPATFULL on STN

ACCESSION NUMBER: 2004:38590 USPATFULL Full-text

TITLE: Identification of essential genes in microorganisms

INVENTOR(S): Wang, Liangsu, San Diego, CA, UNITED STATES
Zamudio, Carlos, La Jolla, CA, UNITED STATES
Malone, Cheryl, Santee, CA, UNITED STATES
Haselbeck, Robert, San Diego, CA, UNITED STATES
Ohlsen, kari L., San Diego, CA, UNITED STATES
Zyskind, Judith W., La Jolla, CA, UNITED STATES

Wall, Daniel, San Diego, CA, UNITED STATES
Trawick, John D., La Mesa, CA, UNITED STATES
Carr, Grant J., Escondido, CA, UNITED STATES
Yamamoto, Robert, San Diego, CA, UNITED STATES
Forsyth, R. Allyn, San Diego, CA, UNITED STATES
Xu, H. Howard, San Diego, CA, UNITED STATES

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2004029129	A1	20040212
APPLICATION INFO.:	US 2002-282122	A1	20021025 (10)

	NUMBER	DATE
PRIORITY INFORMATION:	WO 2002-US9107	20020321
	US 2002-362699P	20020306 (60)
	US 2001-342923P	20011025 (60)
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	APPLICATION	
LEGAL REPRESENTATIVE:	KNOBBE MARTENS OLSON & BEAR LLP, 2040 MAIN STREET, FOURTEENTH FLOOR, IRVINE, CA, 92614	
NUMBER OF CLAIMS:	106	
EXEMPLARY CLAIM:	1	
NUMBER OF DRAWINGS:	22 Drawing Page(s)	
LINE COUNT:	18605	
CAS INDEXING IS AVAILABLE FOR THIS PATENT.		

AB The sequences of antisense nucleic acids which inhibit the proliferation of prokaryotes are disclosed. Cell-based assays which employ the antisense nucleic acids to identify and develop antibiotics are also disclosed. The antisense nucleic acids can also be used to identify proteins required for proliferation, express these proteins or portions thereof, obtain antibodies capable of specifically binding to the expressed proteins, and to use those expressed proteins as a screen to isolate candidate molecules for rational drug discovery programs. The nucleic acids can also be used to screen for homologous nucleic acids that are required for proliferation in cells other than *Staphylococcus aureus*, *Salmonella typhimurium*, *Klebsiella pneumoniae*, and *Pseudomonas aeruginosa*. The nucleic acids of the present invention can also be used in various assay systems to screen for proliferation required genes in other organisms.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

IT 437954-61-5

(amino acid sequence; identification of essential genes in prokaryotes and use of their antisense constructs in antibiotic screening)

L6 ANSWER 28 OF 31 USPATFULL on STN

ACCESSION NUMBER: 2004:250212 USPATFULL Full-text
TITLE: Nucleic acid and amino acid sequences relating to
Streptococcus pneumoniae for diagnostics and
therapeutics
INVENTOR(S): Doucette-Stamm, Lynn A., Framingham, MA, United States
Bush, David, Somerville, MA, United States
PATENT ASSIGNEE(S): Genome Therapeutics Corporation, Waltham, MA, United
States (U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 6800744	B1	20041005
APPLICATION INFO.:	US 1998-107433		19980630 (9)

	NUMBER	DATE
	-----	-----
PRIORITY INFORMATION:	US 1998-85131P	19980512 (60)
	US 1997-51553P	19970702 (60)
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	GRANTED	
PRIMARY EXAMINER:	Brusca, John S.	
ASSISTANT EXAMINER:	Zhou, Shubo "Joe "	
LEGAL REPRESENTATIVE:	Genome Therapeutics Corporation	
NUMBER OF CLAIMS:	14	
EXEMPLARY CLAIM:	1	
NUMBER OF DRAWINGS:	0 Drawing Figure(s); 0 Drawing Page(s)	
LINE COUNT:	11545	

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

AB The invention provides isolated polypeptide and nucleic acid sequences derived from Streptococcus pneumonia that are useful in diagnosis and therapy of pathological conditions; antibodies against the polypeptides; and methods for the production of the polypeptides. The invention also provides methods for the detection, prevention and treatment of pathological conditions resulting from bacterial infection.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

IT 660056-86-0

(amino acid sequence; nucleic acid and amino acid sequences relating to Streptococcus pneumoniae for diagnostics and therapeutics)

L6 ANSWER 29 OF 31 USPATFULL on STN

ACCESSION NUMBER: 2004:59932 USPATFULL Full-text
 TITLE: Staphylococcus epidermidis nucleic acids and proteins
 INVENTOR(S): Kimmerly, William John, Encinitas, CA, United States
 PATENT ASSIGNEE(S): SmithKline Beecham Corporation, Philadelphia, PA, United States (U.S. corporation)

	NUMBER	KIND	DATE
	-----	-----	-----
PATENT INFORMATION:	US 6703492	B1	20040309
APPLICATION INFO.:	US 2000-710279		20001109 (9)

	NUMBER	DATE
	-----	-----
PRIORITY INFORMATION:	US 1999-164258P	19991109 (60)
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	GRANTED	
PRIMARY EXAMINER:	Brusca, John S.	
ASSISTANT EXAMINER:	Zhou, Shubo "Joe"	
LEGAL REPRESENTATIVE:	Conger, Michael M.	
NUMBER OF CLAIMS:	5	
EXEMPLARY CLAIM:	1	
NUMBER OF DRAWINGS:	0 Drawing Figure(s); 0 Drawing Page(s)	
LINE COUNT:	1782	

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

AB S epidermidis polypeptides and DNA (RNA) encoding such polypeptides and a procedure for producing such polypeptides by recombinant techniques is disclosed. Also disclosed are methods for utilizing such polypeptides and DNA (RNA) for the treatment of infection, particularly infections arising from S epidermidis. Antagonists against the function of such polypeptides and their use as therapeutics to treat infection are also disclosed. Also disclosed are diagnostic assays for detecting diseases related to the

presence of S epidermidis nucleic acid sequences and the polypeptides in a host. Also disclosed are diagnostic assays for detecting polynucleotides and polypeptides related to S epidermidis.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

IT 341040-34-4

(amino acid sequence; Staphylococcus epidermidis nucleic acids and proteins as diagnostic or therapeutic agents)

L6 ANSWER 30 OF 31 USPATFULL on STN

ACCESSION NUMBER: 2003:37641 USPATFULL Full-text

TITLE: Bacterial promoters and methods of use

INVENTOR(S): Haselbeck, Robert, San Diego, CA, UNITED STATES

Wall, Daniel, San Diego, CA, UNITED STATES

Gross, Molly, San Diego, CA, UNITED STATES

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2003027286	A1	20030206
APPLICATION INFO.:	US 2001-32393	A1	20011221 (10)

	NUMBER	DATE
PRIORITY INFORMATION:	US 2000-259434P	20001227 (60)
	US 2000-230335P	20000906 (60)
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	APPLICATION	
LEGAL REPRESENTATIVE:	KNOBBE MARTENS OLSON & BEAR LLP, 620 NEWPORT CENTER DRIVE, SIXTEENTH FLOOR, NEWPORT BEACH, CA, 92660	
NUMBER OF CLAIMS:	135	
EXEMPLARY CLAIM:	1	
NUMBER OF DRAWINGS:	17 Drawing Page(s)	
LINE COUNT:	9146	

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

AB Compositions and methods are disclosed herein that relate to the development of fusion promoters for regulating gene expression in bacteria. Embodiments include fusion promoters comprising one or more operators linked to a promoter that is modified to have altered activity in Gram-positive organisms. Vectors and cells containing these fusion promoters are also described. Other embodiments include, methods of using these fusion promoters to regulate nucleic acid and/or polypeptide expression, methods of using these fusion promoters to identify proliferation-required genes, and methods of using these fusion promoters to identify molecules having potential antibiotic activity.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

IT 437954-61-5

(amino acid sequence; identification of essential genes in prokaryotes and use of their antisense constructs in antibiotic screening)

L6 ANSWER 31 OF 31 USPATFULL on STN

ACCESSION NUMBER: 2002:119586 USPATFULL Full-text

TITLE: Identification of essential genes in prokaryotes

INVENTOR(S): Haselbeck, Robert, San Diego, CA, UNITED STATES

Ohlsen, Kari L., San Diego, CA, UNITED STATES

Zyskind, Judith W., La Jolla, CA, UNITED STATES

Wall, Daniel, San Diego, CA, UNITED STATES

Trawick, John D., La Mesa, CA, UNITED STATES

Carr, Grant J., Escondido, CA, UNITED STATES
Yamamoto, Robert T., San Diego, CA, UNITED STATES
Xu, H. Howard, San Diego, CA, UNITED STATES

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2002061569	A1	20020523
APPLICATION INFO.:	US 2001-815242	A1	20010321 (9)

	NUMBER	DATE
PRIORITY INFORMATION:	US 2000-191078P	20000321 (60)
	US 2000-206848P	20000523 (60)
	US 2000-207727P	20000526 (60)
	US 2000-242578P	20001023 (60)
	US 2000-253625P	20001127 (60)
	US 2000-257931P	20001222 (60)
	US 2001-269308P	20010216 (60)

DOCUMENT TYPE: Utility
FILE SEGMENT: APPLICATION
LEGAL REPRESENTATIVE: KNOBBE MARTENS OLSON & BEAR LLP, 620 NEWPORT CENTER
DRIVE, SIXTEENTH FLOOR, NEWPORT BEACH, CA, 92660
NUMBER OF CLAIMS: 44
EXEMPLARY CLAIM: 1
NUMBER OF DRAWINGS: 4 Drawing Page(s)
LINE COUNT: 30870

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

AB The sequences of antisense nucleic acids which inhibit the proliferation of prokaryotes are disclosed. Cell-based assays which employ the antisense nucleic acids to identify and develop antibiotics are also disclosed. The antisense nucleic acids can also be used to identify proteins required for proliferation, express these proteins or portions thereof, obtain antibodies capable of specifically binding to the expressed proteins, and to use those expressed proteins as a screen to isolate candidate molecules for rational drug discovery programs. The nucleic acids can also be used to screen for homologous nucleic acids that are required for proliferation in cells other than *Staphylococcus aureus*, *Salmonella typhimurium*, *Klebsiella pneumoniae*, and *Pseudomonas aeruginosa*. The nucleic acids of the present invention can also be used in various assay systems to screen for proliferation required genes in other organisms.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

IT 437954-61-5
(amino acid sequence; identification of essential genes in prokaryotes and use of their antisense constructs in antibiotic screening)

FILE 'HOME' ENTERED AT 15:08:17 ON 01 DEC 2005

=>

OM protein - protein search, using sw model

Run on: December 2, 2005, 09:25:07 ; Search time 150.286 Seconds
(without alignments)
23.389 Million cell updates/sec

Title: US-10-789-494B-2
Perfect score: 45
Sequence: 1 NINDFDED 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	45	100.0	8	8	ADU51206	Adu51206 Silkworm
2	45	100.0	35	9	AEB30831	Aeb30831 Spider th
3	45	100.0	151	8	ADU51163	Adu51163 Domestic
4	45	100.0	654	9	ADZ09405	Adz09405 Canine pa
5	40	88.9	520	6	ADA35236	Ada35236 Acinetoba
6	39	86.7	1158	4	ABB67681	Abb67681 Drosophil
7	37	82.2	345	5	ABB54921	Abb54921 Lactococc
8	35	77.8	141	6	ADA35940	Ada35940 Acinetoba

9	35	77.8	213	8	ADN74385	Adn74385	Thale cre
10	35	77.8	648	7	ADC70458	Adc70458	Yeast 648
11	35	77.8	648	7	ADK63506	Adk63506	Disease t
12	35	77.8	705	7	ADC23481	Adc23481	Bacillus
13	35	77.8	719	5	ABB48869	Abb48869	Listeria
14	34	75.6	67	2	AAY31453	Aay31453	A. thalia
15	34	75.6	67	3	AAB25813	Aab25813	AP2 direc
16	34	75.6	210	4	ABB68753	Abb68753	Drosophil
17	34	75.6	263	3	AAB22776	Aab22776	Rhizomuco
18	34	75.6	263	3	AAB03824	Aab03824	Orotidine
19	34	75.6	264	7	ADG88453	Adg88453	Arabidops
20	34	75.6	318	4	AAG90278	Aag90278	C glutami
21	34	75.6	318	9	AEB15201	Aeb15201	C glutami
22	34	75.6	320	8	ADY10133	Ady10133	Plant ful
23	34	75.6	321	5	ABP62800	Abp62800	Protein f
24	34	75.6	321	7	ADJ72210	Adj72210	S roseosp
25	34	75.6	335	8	ADX72594	Adx72594	Plant ful
26	34	75.6	364	3	AAG29465	Aag29465	Arabidops
27	34	75.6	432	3	AAG29464	Aag29464	Arabidops
28	34	75.6	432	8	ADO61535	Ado61535	Transcrip
29	34	75.6	432	8	ADN72147	Adn72147	Thale cre
30	34	75.6	449	8	ADX67154	Adx67154	Plant ful
31	34	75.6	509	5	ABP65696	Abp65696	Bifidobac
32	34	75.6	521	4	AAM78789	Aam78789	Human pro
33	34	75.6	585	5	ABP66038	Abp66038	Bifidobac
34	34	75.6	600	4	ABB64253	Abb64253	Drosophil
35	34	75.6	711	8	ADK16463	Adk16463	Nanoarcha
36	34	75.6	794	6	ADA89694	Ada89694	Staphyloc
37	34	75.6	911	6	ABU43640	Abu43640	Protein e
38	34	75.6	917	4	AAU34107	Aau34107	Staphyloc
39	34	75.6	917	6	ABU15958	Abu15958	Protein e
40	34	75.6	917	9	ADW94884	Adw94884	Prolifera
41	34	75.6	920	4	AAU37402	Aau37402	Staphyloc
42	34	75.6	920	4	AAU37555	Aau37555	Staphyloc
43	34	75.6	920	4	AAU36588	Aau36588	Staphyloc
44	34	75.6	920	6	ABM71269	Abm71269	Staphyloc
45	34	75.6	925	8	ADJ48357	Adj48357	Maize oil

ALIGNMENTS

RESULT 1

ADU51206

ID ADU51206 standard; peptide; 8 AA.

XX

AC ADU51206;

XX

DT 24-FEB-2005 (first entry)

XX

DE Silkworm fibroin-derived fibroblast proliferation peptide 3.

XX

KW vulnerary; cell proliferation; wound healing; cell adhesion; cosmetics;
KW cell culture; fibroin.

XX

OS Bombycoidea.

OS Synthetic.

XX
 PN JP2004339189-A.
 XX
 PD 02-DEC-2004.
 XX
 PF 04-DEC-2003; 2003JP-00406608.
 XX
 PR 28-FEB-2003; 2003JP-00055048.
 XX
 PA (DOKU-) DOKURITSU GYOSEI HOJIN NOGYO SEIBUTSU SH.
 PA (TSUB/) TSUBOUCHI K.
 XX
 DR WPI; 2004-827614/82.
 XX
 PT New peptide having excellent cell growth promoting activity, for use as a
 PT cell growth promoter, cell adhesion agent, wound healing-promoting agent,
 PT cosmetic and cell culture base material.
 XX
 PS Claim 2; Page; 27pp; Japanese.
 XX
 CC The invention relates to a novel peptide having excellent cell growth
 CC promoting activity. The peptide of the invention demonstrates vulnerary
 CC activity and may be utilised as a cell growth promoter, cell adhesion
 CC agent, wound healing-promoting agent or cosmetic and cell culture base
 CC material. The current sequence is that of a silkworm fibroin-derived
 CC fibroblast proliferation peptide of the invention.
 XX
 SQ Sequence 8 AA;

Query Match 100.0%; Score 45; DB 8; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NINDFDED 8
 |||||
 Db 1 NINDFDED 8

RESULT 2

AEB30831

ID AEB30831 standard; peptide; 35 AA.

XX

AC AEB30831;

XX

DT 06-OCT-2005 (first entry)

XX

DE Spider thread peptide #2.

XX

KW Silk; spider thread protein.

XX

OS Bombyx mori.

XX

PN WO2005068495-A1.

XX

PD 28-JUL-2005.

XX

PF 12-JAN-2005; 2005WO-JP000619.

XX
 PR 13-JAN-2004; 2004JP-00005489.
 XX
 PA (TORA) TORAY IND INC.
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Hiramatsu S, Moriyama H, Asaoka R, Morita K, Tanaka T, Yamada K;
 PI Obrien JP, Fahnestock SR;
 XX
 DR WPI; 2005-522809/53.
 XX
 PT Silk thread useful for producing textile fabric and in aeronautical
 PT navigation, space exploration, has spider thread protein, produced by
 PT transducing gene encoding spider thread protein to silkworm having
 PT fibroin H-chain gene.
 XX
 PS Claim 17; SEQ ID NO 4; 48pp; Japanese.
 XX
 CC The invention relates to a silk thread comprising a spider thread
 CC protein, produced by a transducing gene encoding spider thread protein in
 CC a silkworm having a fibroin H-chain gene, without damaging the silkworm
 CC fibroin H-chain gene. The invention also relates to producing silk thread
 CC involving producing a transgenic silkworm and extracting silk thread from
 CC the transgenic silkworm. The silk thread is useful for producing a
 CC textile fabric and also useful in aeronautical navigation, space
 CC exploration, to produce clothing, towrope and medical thread, etc. The
 CC silk thread has high strength and elongation property. This sequence
 CC represents a spider thread peptide of the invention.
 XX
 SQ Sequence 35 AA;

Query Match 100.0%; Score 45; DB 9; Length 35;
 Best Local Similarity 100.0%; Pred. No. 0.94;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NINDFDED 8
 |||||
 Db 22 NINDFDED 29

RESULT 5
 ADA35236
 ID ADA35236 standard; protein; 520 AA.
 XX
 AC ADA35236;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Acinetobacter baumannii protein #2397.
 XX
 KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 KW plant biocontrol agent.
 XX
 OS Acinetobacter baumannii.
 XX
 PN US6562958-B1.
 XX

PD 13-MAY-2003.
 XX
 PF 04-JUN-1999; 99US-00328352.
 XX
 PR 09-JUN-1998; 98US-0088701P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton G, Bush D;
 XX
 DR WPI; 2003-576092/54.
 DR N-PSDB; ADA31110.
 XX
 PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 XX
 PS Example; SEQ ID NO 6523; 328pp; English.
 XX
 CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.
 XX
 SQ Sequence 520 AA;

Query Match 88.9%; Score 40; DB 6; Length 520;
 Best Local Similarity 87.5%; Pred. No. 1.3e+02;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NINDFDED 8
 :|||||||
 Db 323 DINDFDED 330

Search completed: December 2, 2005, 09:38:22
 Job time : 154.286 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2005, 09:24:51 ; Search time 22.8571 Seconds
 (without alignments)
 28.936 Million cell updates/sec

Title: US-10-789-494B-2
 Perfect score: 45

Sequence: 1 NINDFDED 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
- 2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
- 3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
- 4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
- 5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	88.9	520	2	US-09-328-352-6523	Sequence 6523, Ap
2	35	77.8	141	2	US-09-328-352-7227	Sequence 7227, Ap
3	35	77.8	292	2	US-09-248-796A-18458	Sequence 18458, A
4	35	77.8	479	2	US-09-248-796A-20464	Sequence 20464, A
5	35	77.8	1009	2	US-09-248-796A-15100	Sequence 15100, A
6	34	75.6	9	2	US-09-629-732-3	Sequence 3, Appli
7	34	75.6	67	1	US-08-700-152A-1	Sequence 1, Appli
8	34	75.6	67	2	US-08-912-272-4	Sequence 4, Appli
9	34	75.6	67	2	US-09-026-039-4	Sequence 4, Appli
10	34	75.6	67	2	US-08-879-827A-4	Sequence 4, Appli
11	34	75.6	182	2	US-09-270-767-34346	Sequence 34346, A
12	34	75.6	182	2	US-09-270-767-49563	Sequence 49563, A
13	34	75.6	432	1	US-08-700-152A-4	Sequence 4, Appli
14	34	75.6	1076	2	US-09-976-594-889	Sequence 889, App
15	33	73.3	167	2	US-09-252-991A-21860	Sequence 21860, A
16	33	73.3	231	2	US-09-134-000C-4214	Sequence 4214, Ap
17	33	73.3	318	2	US-09-248-796A-20696	Sequence 20696, A
18	33	73.3	422	1	US-08-680-726A-68	Sequence 68, Appl
19	33	73.3	422	2	US-09-092-409-68	Sequence 68, Appl
20	33	73.3	690	2	US-09-388-743-6	Sequence 6, Appli
21	33	73.3	690	2	US-10-044-543-6	Sequence 6, Appli
22	33	73.3	738	2	US-09-248-796A-20896	Sequence 20896, A
23	32	71.1	65	1	US-08-227-536-8	Sequence 8, Appli
24	32	71.1	65	4	PCT-US95-04682-8	Sequence 8, Appli

25	32	71.1	128	2	US-09-252-991A-29019	Sequence 29019, A
26	32	71.1	130	2	US-09-248-796A-23298	Sequence 23298, A
27	32	71.1	155	2	US-09-540-236-3682	Sequence 3682, Ap
28	32	71.1	180	2	US-09-248-796A-27668	Sequence 27668, A
29	32	71.1	545	2	US-09-248-796A-24484	Sequence 24484, A
30	32	71.1	591	2	US-09-248-796A-14242	Sequence 14242, A
31	32	71.1	617	2	US-09-248-796A-26692	Sequence 26692, A
32	32	71.1	779	2	US-09-749-601A-12	Sequence 12, Appl
33	32	71.1	805	2	US-09-425-335-6	Sequence 6, Appli
34	32	71.1	868	2	US-09-248-796A-16660	Sequence 16660, A
35	32	71.1	1217	2	US-09-949-016-7454	Sequence 7454, Ap
36	32	71.1	1417	1	US-08-559-303B-78	Sequence 78, Appl
37	32	71.1	1417	2	US-08-781-891-78	Sequence 78, Appl
38	32	71.1	1417	2	US-09-175-828-78	Sequence 78, Appl
39	32	71.1	1417	2	US-09-618-166-78	Sequence 78, Appl
40	32	71.1	1417	2	US-09-753-143-78	Sequence 78, Appl
41	32	71.1	1507	6	5268270-2	Patent No. 5268270
42	31	68.9	26	2	US-09-962-756-414	Sequence 414, App
43	31	68.9	207	2	US-09-270-767-32417	Sequence 32417, A
44	31	68.9	212	2	US-09-710-279-800	Sequence 800, App
45	31	68.9	220	1	US-08-840-683-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1

US-09-328-352-6523

; Sequence 6523, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 6523

; LENGTH: 520

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-6523

Query Match 88.9%; Score 40; DB 2; Length 520;

Best Local Similarity 87.5%; Pred. No. 29;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 NINDFDED 8
          :|||||
Db     323 DINDFDED 330

```

RESULT 6

US-09-629-732-3

```
; Sequence 3, Application US/09629732
; Patent No. 6631329
; GENERAL INFORMATION:
; APPLICANT: Yale University
; APPLICANT: STEITZ, Thomas A.
; APPLICANT: WANG, Jimin
; APPLICANT: SILVIAN, Laura F.
; TITLE OF INVENTION: Use of the Crystal Structure of Staphylococcus Aureus
Isoleucyl-tRNA
; TITLE OF INVENTION: Synthetase in Antibiotic Design
; FILE REFERENCE: 44574-5075-US
; CURRENT APPLICATION NUMBER: US/09/629,732
; CURRENT FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/146,176
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Probe Z
US-09-629-732-3
```

```
Query Match          75.6%; Score 34; DB 2; Length 9;
Best Local Similarity 75.0%; Pred. No. 4.6e+05;
Matches      6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 NINDFDED 8
        |||||: |
Db      2 NINDFNPD 9
```

```
Search completed: December 2, 2005, 09:33:50
Job time : 23.8571 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

```
Run on:      December 2, 2005, 09:28:17 ; Search time 125.714 Seconds
              (without alignments)
              26.589 Million cell updates/sec
```

```
Title:      US-10-789-494B-2
Perfect score: 45
Sequence:    1 NINDFDED 8
```

```
Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5
```

```
Searched:    1867569 seqs, 417829326 residues
```

```
Total number of hits satisfying chosen parameters: 1867569
```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	45	100.0	8	5	US-10-789-494B-2	Sequence 2, Appli
2	45	100.0	151	5	US-10-789-494B-9	Sequence 9, Appli
3	39	86.7	1158	6	US-11-097-143-29835	Sequence 29835, A
4	38	84.4	297	4	US-10-437-963-181331	Sequence 181331, A
5	38	84.4	441	4	US-10-156-761-13669	Sequence 13669, A
6	38	84.4	919	4	US-10-437-963-182815	Sequence 182815, A
7	38	84.4	1301	4	US-10-437-963-182743	Sequence 182743, A
8	38	84.4	1888	4	US-10-437-963-142571	Sequence 142571, A
9	37	82.2	1974	4	US-10-437-963-184754	Sequence 184754, A
10	36	80.0	143	4	US-10-425-115-333223	Sequence 333223, A
11	36	80.0	343	4	US-10-424-599-275987	Sequence 275987, A
12	36	80.0	373	4	US-10-767-701-45662	Sequence 45662, A
13	36	80.0	373	4	US-10-425-115-290243	Sequence 290243, A
14	36	80.0	373	4	US-10-425-115-290246	Sequence 290246, A
15	36	80.0	548	4	US-10-425-115-333222	Sequence 333222, A
16	35	77.8	185	4	US-10-424-599-203609	Sequence 203609, A
17	35	77.8	705	5	US-10-504-543-2	Sequence 2, Appli
18	34	75.6	183	4	US-10-437-963-202547	Sequence 202547, A
19	34	75.6	210	6	US-11-097-143-33051	Sequence 33051, A
20	34	75.6	264	4	US-10-059-911-24	Sequence 24, Appl
21	34	75.6	287	4	US-10-425-115-205811	Sequence 205811, A
22	34	75.6	318	3	US-09-738-626-4032	Sequence 4032, Ap
23	34	75.6	318	6	US-11-006-098-116	Sequence 116, App
24	34	75.6	320	4	US-10-425-114-65948	Sequence 65948, A
25	34	75.6	320	4	US-10-425-115-295750	Sequence 295750, A
26	34	75.6	321	5	US-10-211-028-85	Sequence 85, Appl
27	34	75.6	335	4	US-10-425-114-41960	Sequence 41960, A
28	34	75.6	449	4	US-10-425-114-37997	Sequence 37997, A
29	34	75.6	472	4	US-10-156-761-9469	Sequence 9469, Ap
30	34	75.6	515	4	US-10-437-963-181539	Sequence 181539, A
31	34	75.6	600	6	US-11-097-143-19551	Sequence 19551, A
32	34	75.6	911	4	US-10-282-122A-71564	Sequence 71564, A

33	34	75.6	917	3	US-09-815-242-5603	Sequence 5603, Ap
34	34	75.6	917	4	US-10-282-122A-43882	Sequence 43882, A
35	34	75.6	917	5	US-10-857-625-824	Sequence 824, App
36	34	75.6	920	3	US-09-815-242-12181	Sequence 12181, A
37	34	75.6	920	3	US-09-815-242-12995	Sequence 12995, A
38	34	75.6	920	3	US-09-815-242-13148	Sequence 13148, A
39	34	75.6	925	4	US-10-389-566-361	Sequence 361, App
40	34	75.6	1023	5	US-10-450-763-53242	Sequence 53242, A
41	34	75.6	1076	4	US-10-275-595A-26	Sequence 26, Appl
42	34	75.6	1171	4	US-10-312-042-8	Sequence 8, Appli
43	34	75.6	1288	5	US-10-732-923-8591	Sequence 8591, Ap
44	34	75.6	4226	5	US-10-732-923-22586	Sequence 22586, A
45	34	75.6	4226	5	US-10-732-923-22707	Sequence 22707, A

ALIGNMENTS

RESULT 1

US-10-789-494B-2

```
; Sequence 2, Application US/10789494B
; Publication No. US20050143296A1
; GENERAL INFORMATION:
; APPLICANT: TSUBOUCHI, Kozo
; APPLICANT: YAMADA, Hiromi
; TITLE OF INVENTION: EXTRACTION AND UTILIZATION OF CELL
; TITLE OF INVENTION: GROWTH-PROMOTING PEPTIDES FROM SILK PROTEIN
; FILE REFERENCE: OPS 635
; CURRENT APPLICATION NUMBER: US/10/789,494B
; CURRENT FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: JP 2003-55048
; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 2
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Bombyx mori
US-10-789-494B-2
```

```
Query Match          100.0%;  Score 45;  DB 5;  Length 8;
Best Local Similarity 100.0%;  Pred. No. 1.7e+06;
Matches      8;  Conservative    0;  Mismatches    0;  Indels      0;  Gaps      0;
```

```
QY      1 NINDFDED 8
          |||||
Db      1 NINDFDED 8
```

RESULT 3

US-11-097-143-29835

```
; Sequence 29835, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
```

```

; TITLE OF INVENTION:  ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION:  DROSOPHILA GENES.
; FILE REFERENCE:  CL000728
; CURRENT APPLICATION NUMBER:  US/11/097,143
; CURRENT FILING DATE:  2005-04-04
; PRIOR APPLICATION NUMBER:  60/157,832
; PRIOR FILING DATE:  1999-10-05
; PRIOR APPLICATION NUMBER:  60/160,191
; PRIOR FILING DATE:  1999-10-19
; PRIOR APPLICATION NUMBER:  60/161,932
; PRIOR FILING DATE:  1999-10-28
; PRIOR APPLICATION NUMBER:  60/164,769
; PRIOR FILING DATE:  1999-11-12
; PRIOR APPLICATION NUMBER:  60/173,383
; PRIOR FILING DATE:  1999-12-28
; PRIOR APPLICATION NUMBER:  60/175,693
; PRIOR FILING DATE:  2000-01-12
; PRIOR APPLICATION NUMBER:  60/184,831
; PRIOR FILING DATE:  2000-02-24
; PRIOR APPLICATION NUMBER:  60/191,637
; PRIOR FILING DATE:  2000-03-23
; NUMBER OF SEQ ID NOS:  43008
; SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 29835
;   LENGTH:  1158
;   TYPE:  PRT
;   ORGANISM:  DROSOPHILA
US-11-097-143-29835

```

```

Query Match          86.7%;  Score 39;  DB 6;  Length 1158;
Best Local Similarity  75.0%;  Pred. No. 3.9e+02;
Matches      6;  Conservative      2;  Mismatches      0;  Indels      0;  Gaps      0;

```

```

Qy      1 NINDFDED 8
        |::|||||
Db      613 NVSDFDED 620

```

```

Search completed: December  2, 2005, 09:55:39
Job time : 126.714 secs

```

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005  Compugen Ltd.

```

OM protein - protein search, using sw model

```

Run on:      December  2, 2005, 09:33:57 ; Search time 6.85714 Seconds
              (without alignments)
              5.586 Million cell updates/sec

```

```

Title:      US-10-789-494B-2
Perfect score:  45
Sequence:    1 NINDFDED 8

```

```

Scoring table:  BLOSUM62
                 Gapop 10.0 , Gapext 0.5

```

Searched: 26661 seqs, 4788334 residues

Total number of hits satisfying chosen parameters: 26661

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA_New:*

- 1: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	35	77.8	648	6	US-10-501-039-6	Sequence 6, Appli
2	34	75.6	794	6	US-10-485-517-355	Sequence 355, App
3	32	71.1	257	6	US-10-485-517-388	Sequence 388, App
4	32	71.1	1565	6	US-10-467-657-2704	Sequence 2704, Ap
5	31	68.9	212	6	US-10-793-626-800	Sequence 800, App
6	30	66.7	313	6	US-10-793-626-1758	Sequence 1758, Ap
7	30	66.7	948	6	US-10-485-517-131	Sequence 131, App
8	30	66.7	1992	7	US-11-013-759-3	Sequence 3, Appli
9	30	66.7	1992	7	US-11-013-759-13	Sequence 13, Appl
10	30	66.7	2047	7	US-11-013-759-4	Sequence 4, Appli
11	30	66.7	2047	7	US-11-013-759-7	Sequence 7, Appli
12	29	64.4	106	6	US-10-793-626-794	Sequence 794, App
13	29	64.4	106	6	US-10-793-626-2140	Sequence 2140, Ap
14	29	64.4	221	6	US-10-793-626-2778	Sequence 2778, Ap
15	29	64.4	655	6	US-10-793-626-1052	Sequence 1052, Ap
16	29	64.4	655	6	US-10-793-626-1400	Sequence 1400, Ap
17	29	64.4	1151	6	US-10-793-626-2448	Sequence 2448, Ap
18	29	64.4	1663	6	US-10-982-545-6	Sequence 6, Appli
19	28	62.2	314	6	US-10-793-626-3310	Sequence 3310, Ap
20	28	62.2	320	6	US-10-467-657-1360	Sequence 1360, Ap
21	28	62.2	381	6	US-10-793-626-3056	Sequence 3056, Ap
22	28	62.2	446	6	US-10-793-626-1836	Sequence 1836, Ap
23	28	62.2	512	6	US-10-821-234-1032	Sequence 1032, Ap
24	28	62.2	566	6	US-10-467-657-4020	Sequence 4020, Ap
25	28	62.2	600	6	US-10-467-657-4866	Sequence 4866, Ap
26	28	62.2	685	6	US-10-131-826A-88	Sequence 88, Appl

27	28	62.2	685	7	US-11-078-735-19	Sequence 19, Appl
28	28	62.2	945	6	US-10-131-826A-146	Sequence 146, App
29	28	62.2	1616	6	US-10-821-234-1497	Sequence 1497, Ap
30	27	60.0	93	6	US-10-467-657-8731	Sequence 8731, Ap
31	27	60.0	113	6	US-10-845-413-297	Sequence 297, App
32	27	60.0	113	6	US-10-845-413-299	Sequence 299, App
33	27	60.0	140	6	US-10-467-657-2486	Sequence 2486, Ap
34	27	60.0	184	7	US-11-074-176-16	Sequence 16, Appl
35	27	60.0	195	7	US-11-038-284-26	Sequence 26, Appl
36	27	60.0	230	6	US-10-467-657-952	Sequence 952, App
37	27	60.0	236	6	US-10-467-657-5368	Sequence 5368, Ap
38	27	60.0	296	6	US-10-793-626-1674	Sequence 1674, Ap
39	27	60.0	364	6	US-10-467-657-2822	Sequence 2822, Ap
40	27	60.0	435	6	US-10-467-657-318	Sequence 318, App
41	27	60.0	522	6	US-10-793-626-604	Sequence 604, App
42	27	60.0	523	6	US-10-467-657-5392	Sequence 5392, Ap
43	27	60.0	551	6	US-10-793-626-1668	Sequence 1668, Ap
44	27	60.0	611	6	US-10-467-657-4656	Sequence 4656, Ap
45	27	60.0	617	6	US-10-982-545-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-10-501-039-6

; Sequence 6, Application US/10501039

; Publication No. US20050244822A1

; GENERAL INFORMATION:

; APPLICANT: Tetsuro Kokubo, Masahiro Shirakawa, and Jeremy Robin Howard Tame

; TITLE OF INVENTION: Method of monitoring gene expression

; FILE REFERENCE: 4439-4023

; CURRENT APPLICATION NUMBER: US/10/501,039

; CURRENT FILING DATE: 2004-07-08

; PRIOR APPLICATION NUMBER: JP P2002-002396

; PRIOR FILING DATE: 2002-01-09

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 648

; TYPE: PRT

; ORGANISM: *Saccharomyces cerevisiae*

US-10-501-039-6

Query Match 77.8%; Score 35; DB 6; Length 648;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      3 NDFDED 8
          |||||
Db     507 NDFDED 512

```

Search completed: December 2, 2005, 09:56:15

Job time : 7.85714 secs

OM protein - protein search, using sw model

Run on: December 2, 2005, 09:38:38 ; Search time 26.8571 Seconds
(without alignments)
28.660 Million cell updates/sec

Title: US-10-789-494B-2
Perfect score: 45
Sequence: 1 NINDFDED 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	38	84.4	395	2	AD3354	hypothetical cytos
2	38	84.4	446	2	T35094	hypothetical prote
3	38	84.4	576	1	B70558	probable ABC trans
4	37	82.2	232	2	JQ1199	replication protei
5	37	82.2	233	2	S15954	repB protein - Lac
6	37	82.2	345	2	G86821	S-adenosylmethioni
7	35	77.8	117	2	G97840	hypothetical prote
8	35	77.8	134	2	B35119	4-carboxymuconolac
9	35	77.8	164	1	RNVZ19	DNA-directed RNA p
10	35	77.8	164	2	T28547	hypothetical prote
11	35	77.8	164	2	C72164	A6R protein - vari
12	35	77.8	164	2	F36848	A5R protein - vari
13	35	77.8	213	2	F84581	copia-like retroel

14	35	77.8	435	2	T30114	hypothetical prote
15	35	77.8	468	2	S49391	GltX protein - Myc
16	35	77.8	550	2	S55118	probable membrane
17	35	77.8	648	2	S56783	hypothetical prote
18	35	77.8	719	2	AI1212	TN916 ORF15 homolo
19	35	77.8	729	2	T52187	probable transposa
20	35	77.8	851	2	A86200	hypothetical prote
21	34	75.6	97	2	E97266	glu-tRNA amidotran
22	34	75.6	132	2	D72151	B12L protein - var
23	34	75.6	152	2	T28445	hypothetical prote
24	34	75.6	153	2	G36837	D7L protein - vari
25	34	75.6	295	2	S61039	hypothetical prote
26	34	75.6	319	2	AF2199	hypothetical prote
27	34	75.6	340	2	H81346	hypothetical prote
28	34	75.6	386	2	D42528	B23R protein - vac
29	34	75.6	432	2	A85436	APETALA2 protein [
30	34	75.6	469	2	T34645	hypothetical prote
31	34	75.6	634	1	WZVZA8	74K HindIII-C prot
32	34	75.6	634	2	E42503	C9L protein - vacc
33	34	75.6	917	2	S40178	isoleucine-tRNA li
34	34	75.6	917	2	D89891	Ile-tRNA synthetas
35	34	75.6	1171	2	T17454	diaphanous-related
36	34	75.6	1288	2	T37528	probable snf2 fami
37	33	73.3	105	2	H86863	hypothetical prote
38	33	73.3	255	2	T49972	hypothetical prote
39	33	73.3	271	2	T24965	hypothetical prote
40	33	73.3	316	2	T33180	hypothetical prote
41	33	73.3	340	2	F82468	hypothetical prote
42	33	73.3	372	2	C81263	probable integral
43	33	73.3	407	2	C45600	asparagine-rich bl
44	33	73.3	449	2	T44643	galactosyl transfe
45	33	73.3	463	2	T28748	hypothetical prote

ALIGNMENTS

RESULT 1

AD3354

hypothetical cytosolic protein BMEI0818 [imported] - *Brucella melitensis* (strain 16M)

C;Species: *Brucella melitensis*

C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C;Accession: AD3354

R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.;

Ivanova, N.; Anderson, I.; Bhattacharyya, A.; Lykidis, A.; Reznik, G.;

Jablonski, L.; Larsen, N.; D'Souza, M.; Bernal, A.; Mazur, M.; Goltsman, E.;

Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesson, J.J.; Haselkorn,

R.; Kyrpides, N.; Overbeek, R.

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*.

A;Reference number: AD3252; PMID:11756688

A;Accession: AD3354

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-395 <KUR>

A;Cross-references: UNIPROT:Q8YHI1; UNIPARC:UPI0000057E1C; GB:AE008917;
PIDN:AAL51999.1; PID:g17982762; GSPDB:GN00190
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI0818
A;Map position: I

Query Match 84.4%; Score 38; DB 2; Length 395;
Best Local Similarity 87.5%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NINDFDED 8
||| ||
Db 19 NINDFTED 26

Search completed: December 2, 2005, 09:57:10
Job time : 29.8571 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2005, 09:24:01 ; Search time 166.857 Seconds
(without alignments)
33.827 Million cell updates/sec

Title: US-10-789-494B-2
Perfect score: 45
Sequence: 1 NINDFDED 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

%

Result No.	Score	Query Match	Length	DB	ID	Description
1	45	100.0	178	1	FIBH_BOMMA	Q99050 bombyx mand
2	45	100.0	5263	1	FIBH_BOMMO	P05790 bombyx mori
3	39	86.7	186	2	Q899G8_CLOTE	Q899g8 clostridium
4	39	86.7	867	2	Q8SZE7_DROME	Q8sze7 drosophila
5	39	86.7	941	2	Q9VXA2_DROME	Q9vxa2 drosophila
6	38	84.4	133	2	Q7RQS0_PLAYO	Q7rqs0 plasmodium
7	38	84.4	382	2	Q512U2_ENTHI	Q512u2 entamoeba h
8	38	84.4	395	2	Q57CX2_BRUAB	Q57cx2 brucella ab
9	38	84.4	395	2	Q8G0D1_BRUSU	Q8g0d1 brucella su
10	38	84.4	395	2	Q8YHI1_BRUME	Q8yhi1 brucella me
11	38	84.4	441	2	Q82AB7_STRAW	Q82ab7 streptomyce
12	38	84.4	446	2	Q9S2S5_STRCO	Q9s2s5 streptomyce
13	38	84.4	448	2	Q5BA42_EMENI	Q5ba42 aspergillus
14	38	84.4	481	2	Q69HQ9_CIOIN	Q69hq9 ciona intes
15	38	84.4	576	2	O06137_MYCTU	O06137 mycobacteri
16	38	84.4	576	2	Q7TZV7_MYCBO	Q7tzv7 mycobacteri
17	38	84.4	585	2	Q50PU4_ENTHI	Q50pu4 entamoeba h
18	38	84.4	1110	2	Q756I5_ASHGO	Q756i5 ashbya goss
19	38	84.4	1474	2	Q6F2F8_ORYSA	Q6f2f8 oryza sativ
20	38	84.4	2378	2	Q8I3U0_PLAF7	Q8i3u0 plasmodium
21	37	82.2	71	2	Q775Z3_CAMPS	Q775z3 camelpox vi
22	37	82.2	71	2	Q8V2X5_CAMPM	Q8v2x5 camelpox vi
23	37	82.2	232	2	Q04138_9LACT	Q04138 lactococcus
24	37	82.2	232	2	Q48821_LACPL	Q48821 lactobacill
25	37	82.2	345	1	QUEA_LACLA	Q9cfa6 lactococcus
26	37	82.2	402	2	Q8ET06_OCEIH	Q8et06 oceanobacil
27	37	82.2	418	2	Q26662_STRPU	Q26662 strongyloce
28	37	82.2	842	2	Q6FR44_CANGA	Q6fr44 candida gla
29	37	82.2	923	2	Q747X9_GEOSL	Q747x9 geobacter s
30	37	82.2	1416	1	BLM_MOUSE	O88700 mus musculu
31	36	80.0	309	2	Q5FJN9_LACAC	Q5fjn9 lactobacill
32	36	80.0	323	2	Q6LRT6_PHOPR	Q6lrt6 photobacter
33	36	80.0	500	2	Q54L47_DICDI	Q54l47 dictyosteli
34	36	80.0	515	2	Q65MB6_BACLD	Q65mb6 bacillus li
35	36	80.0	554	2	Q73MY4_TREDE	Q73my4 treponema d
36	36	80.0	557	2	Q51AZ1_ENTHI	Q51az1 entamoeba h
37	36	80.0	687	2	Q4YVU0_PLABE	Q4yvu0 plasmodium
38	36	80.0	876	2	Q75JU2_DICDI	Q75ju2 dictyosteli
39	36	80.0	895	2	Q4T021_TETNG	Q4t021 tetraodon n
40	36	80.0	979	2	Q4HMP3_CAMLA	Q4hmp3 campylobact
41	36	80.0	1445	2	Q5CPT3_CRYPV	Q5cpt3 cryptospori
42	35	77.8	117	2	Q92GJ6_RICCN	Q92gj6 rickettsia
43	35	77.8	133	2	Q4X7A0_PLACH	Q4x7a0 plasmodium
44	35	77.8	133	2	Q4YWR1_PLABE	Q4ywr1 plasmodium
45	35	77.8	134	1	DC4C_ACIAD	P20370 acinetobact

ALIGNMENTS

RESULT 1

FIBH_BOMMA

ID FIBH_BOMMA STANDARD; PRT; 178 AA.

AC Q99050;

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Fibroin heavy chain precursor (Fib-H) (H-fibroin) (Fragment).
 GN Name=FIBH;
 OS Bombyx mandarina (Wild silk moth) (Wild silkworm).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
 OC Bombycidae; Bombyx.
 OX NCBI_TaxID=7092;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Posterior silk gland;
 RA Kusuda J., Tazima Y., Onimaru K., Ninaki O., Suzuki Y.;
 RT "The sequence around the 5' end of the fibroin gene from the wild
 RT silkworm, Bombyx mandarina, and comparison with that of the
 RT domesticated species, B. mori.";
 RL Mol. Gen. Genet. 203:359-364(1986).
 CC -!- FUNCTION: Core component of the silk filament; a strong, insoluble
 CC and chemically inert fiber.
 CC -!- SUBUNIT: Silk fibroin elementary unit consists in a disulfide-
 CC linked heavy and light chain and a p25 glycoprotein in molar
 CC ratios of 6:6:1. This results in a complex of approximately 2.3
 CC MDa.
 CC -!- TISSUE SPECIFICITY: Produced exclusively in the posterior (PSG)
 CC section of silk glands, which are essentially modified salivary
 CC glands.
 CC -!- DOMAIN: Composed of antiparallel beta sheets. The strands of the
 CC beta sheets run parallel to the fiber axis. Long stretches of silk
 CC fibroin are composed of microcrystalline arrays of (-Gly-Ser-Gly-
 CC Ala-Gly-Ala-)n interrupted by regions containing bulkier residues.
 CC The fiber is composed of microcrystalline arrays alternating with
 CC amorphous regions.
 CC -!- PTM: The interchain disulfide bridge is essential for the
 CC intracellular transport and secretion of fibroin.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR EMBL; X03973; CAA27612.1; -; Genomic_DNA.
 KW Repeat; Signal; Silk.
 FT SIGNAL 1 21 Potential.
 FT CHAIN 22 >178 Fibroin heavy chain.
 FT REGION 149 >178 Highly repetitive.
 FT CONFLICT 10 10 C -> V (in Ref. 1; CAA27612).
 FT NON_TER 178 178
 SQ SEQUENCE 178 AA; 18326 MW; 8E15C7E7A9682940 CRC64;

Query Match 100.0%; Score 45; DB 1; Length 178;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NINDFDED 8
 |||||

Db 22 NINDFDED 29

Search completed: December 2, 2005, 09:33:08
Job time : 171.857 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2005, 09:25:07 ; Search time 112.714 Seconds
(without alignments)
23.389 Million cell updates/sec

Title: US-10-789-494B-6
Perfect score: 34
Sequence: 1 DEYVDN 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	34	100.0	6	8	ADU51236	Adu51236 Gut silkw
2	34	100.0	6	8	ADU51210	Adu51210 Silkworm
3	34	100.0	120	8	ADU51171	Adu51171 Gut silkw
4	34	100.0	126	8	ADK48385	Adk48385 Streptoco
5	34	100.0	126	8	ADR94878	Adr94878 Novel S.
6	34	100.0	126	9	AEA58748	Aea58748 Streptoco
7	34	100.0	383	4	AAG82666	Aag82666 S. epider
8	34	100.0	383	6	ABJ19176	Abj19176 Pathogen

9	34	100.0	384	5	ABP38592	Abp38592	Staphyloc
10	34	100.0	384	8	ADS06341	Ads06341	Staphyloc
11	34	100.0	568	3	AAB18230	Aab18230	Plasmodiu
12	34	100.0	2368	4	AAU34139	Aau34139	Staphyloc
13	34	100.0	2368	4	AAU36796	Aau36796	Staphyloc
14	34	100.0	2655	7	ADO59401	Ado59401	Antheraea
15	31	91.2	172	8	ADO61859	Ado61859	Transcrip
16	31	91.2	275	8	ADF93905	Adf93905	Carotene
17	31	91.2	346	4	ABG29210	Abg29210	Novel hum
18	31	91.2	354	2	AAW21994	Aaw21994	Tetracycl
19	31	91.2	554	8	ADI79890	Adi79890	Mouse liv
20	31	91.2	567	3	AAY59507	Aay59507	C. elegan
21	31	91.2	567	3	AAB03667	Aab03667	Nematode
22	31	91.2	567	8	ADN23284	Adn23284	Bacterial
23	31	91.2	582	8	ADQ48555	Adq48555	AcMPNV IE
24	31	91.2	597	8	ADN23283	Adn23283	Bacterial
25	31	91.2	704	3	AAY91091	Aay91091	Caenorhab
26	31	91.2	704	3	AAY59506	Aay59506	C. elegan
27	31	91.2	704	5	ABB90799	Abb90799	Herbicida
28	31	91.2	704	8	ADN23282	Adn23282	Bacterial
29	31	91.2	713	2	AAR99797	Aar99797	Lysine de
30	31	91.2	713	8	ADN18055	Adn18055	Bacterial
31	31	91.2	789	8	ADL70332	Adl70332	Crenarcha
32	31	91.2	983	8	ADX67859	Adx67859	Plant ful
33	31	91.2	1268	8	ADF93901	Adf93901	Carotene
34	30	88.2	99	3	AAG15954	Aag15954	Arabidops
35	30	88.2	152	3	AAG15953	Aag15953	Arabidops
36	30	88.2	153	3	AAG15952	Aag15952	Arabidops
37	30	88.2	209	7	ABO72159	Abo72159	Pseudomon
38	30	88.2	226	6	ADB09735	Adb09735	Alloiococ
39	30	88.2	325	2	AAW47420	Aaw47420	Micrococc
40	30	88.2	467	5	AAE23627	Aae23627	Lactococc
41	30	88.2	468	6	ABU46827	Abu46827	Protein e
42	30	88.2	468	6	ABU46068	Abu46068	Protein e
43	30	88.2	468	8	ADR83959	Adr83959	S. pyogen
44	30	88.2	468	8	ADV87930	Adv87930	Streptoco
45	30	88.2	468	8	ADV79183	Adv79183	Streptoco

ALIGNMENTS

RESULT 1

ADU51236

ID ADU51236 standard; peptide; 6 AA.

XX

AC ADU51236;

XX

DT 24-FEB-2005 (first entry)

XX

DE Gut silkworm fibroin peptide fragment 38.

XX

KW vulnerary; cell proliferation; wound healing; cell adhesion; cosmetics;
KW cell culture; fibroin.

XX

OS Bombycoidea.

XX

PN JP2004339189-A.
 XX
 PD 02-DEC-2004.
 XX
 PF 04-DEC-2003; 2003JP-00406608.
 XX
 PR 28-FEB-2003; 2003JP-00055048.
 XX
 PA (DOKU-) DOKURITSU GYOSEI HOJIN NOGYO SEIBUTSU SH.
 PA (TSUB/) TSUBOUCHI K.
 XX
 DR WPI; 2004-827614/82.
 XX
 PT New peptide having excellent cell growth promoting activity, for use as a
 PT cell growth promoter, cell adhesion agent, wound healing-promoting agent,
 PT cosmetic and cell culture base material.
 XX
 PS Example 3; Page; 27pp; Japanese.
 XX
 CC The invention relates to a novel peptide having excellent cell growth
 CC promoting activity. The peptide of the invention demonstrates vulnerary
 CC activity and may be utilised as a cell growth promoter, cell adhesion
 CC agent, wound healing-promoting agent or cosmetic and cell culture base
 CC material. The current sequence is that of a gut silkworm fibroin peptide
 CC fragment of the invention which is described as being amorphous.
 XX
 SQ Sequence 6 AA;

Query Match 100.0%; Score 34; DB 8; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEYVDN 6
 |||||
 Db 1 DEYVDN 6

RESULT 4

ADK48385

ID ADK48385 standard; protein; 126 AA.
 XX
 AC ADK48385;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Streptococcus pneumoniae protein, Seq ID No 4900.
 XX
 KW Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN US6699703-B1.
 XX
 PD 02-MAR-2004.
 XX
 PF 26-MAY-2000; 2000US-00583110.

XX
 PR 02-JUL-1997; 97US-0051553P.
 PR 12-MAY-1998; 98US-0085131P.
 PR 30-JUN-1998; 98US-00107433.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;
 XX
 DR WPI; 2004-212399/20.
 DR N-PSDB; ADK45724.
 XX
 PT New nucleic acid molecules and polypeptides useful for diagnosing,
 PT preventing and treating pathological conditions resulting from bacterial
 PT infection, e.g. Streptococcus pneumoniae infection, and in drug
 PT screening.
 XX
 PS Disclosure; SEQ ID NO 4900; 301pp; English.
 XX
 CC The invention relates to isolated Streptococcus pneumoniae nucleic acids
 CC and polypeptides. The nucleic acids and proteins are useful for
 CC diagnosing, preventing and treating pathological conditions resulting
 CC from bacterial infection, such as S. pneumoniae infection. These may also
 CC be used for drug screening procedures. The present sequence represents a
 CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence
 CC data for this patent did not appear in the printed specification but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 126 AA;

Query Match 100.0%; Score 34; DB 8; Length 126;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEYVDN 6
 |||||
 Db 81 DEYVDN 86

Search completed: December 2, 2005, 09:38:25
 Job time : 115.714 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2005, 09:24:51 ; Search time 17.1429 Seconds
 (without alignments)
 28.936 Million cell updates/sec

Title: US-10-789-494B-6
 Perfect score: 34
 Sequence: 1 DEYVDN 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34	100.0	126	2	US-09-583-110-4900	Sequence 4900, Ap
2	34	100.0	126	2	US-09-107-433-3513	Sequence 3513, Ap
3	34	100.0	383	2	US-09-710-279-2426	Sequence 2426, Ap
4	34	100.0	384	2	US-09-134-001C-3437	Sequence 3437, Ap
5	31	91.2	354	2	US-08-970-264A-21	Sequence 21, Appl
6	31	91.2	713	1	US-08-849-212-4	Sequence 4, Appli
7	30	88.2	88	2	US-09-270-767-40244	Sequence 40244, A
8	30	88.2	88	2	US-09-270-767-55460	Sequence 55460, A
9	30	88.2	209	2	US-09-252-991A-20905	Sequence 20905, A
10	30	88.2	283	2	US-09-248-796A-19476	Sequence 19476, A
11	30	88.2	325	2	US-09-217-609A-2	Sequence 2, Appli
12	30	88.2	325	2	US-08-873-235B-2	Sequence 2, Appli
13	30	88.2	467	2	US-08-914-375C-57	Sequence 57, Appl
14	30	88.2	471	2	US-09-583-110-3861	Sequence 3861, Ap
15	30	88.2	471	2	US-09-107-433-4791	Sequence 4791, Ap
16	30	88.2	495	2	US-09-107-532A-5715	Sequence 5715, Ap
17	29	85.3	14	2	US-09-200-650E-23	Sequence 23, Appl
18	29	85.3	88	2	US-09-248-796A-21726	Sequence 21726, A
19	29	85.3	127	2	US-09-437-054A-12	Sequence 12, Appl
20	29	85.3	160	2	US-09-134-000C-4511	Sequence 4511, Ap
21	29	85.3	395	2	US-09-248-796A-14134	Sequence 14134, A
22	29	85.3	434	2	US-09-487-558B-146	Sequence 146, App
23	29	85.3	473	2	US-09-134-000C-5440	Sequence 5440, Ap
24	29	85.3	524	2	US-09-248-796A-26474	Sequence 26474, A
25	29	85.3	634	2	US-09-949-016-10571	Sequence 10571, A

26	29	85.3	670	2	US-09-543-681A-5979	Sequence 5979, Ap
27	29	85.3	720	2	US-09-257-799-48	Sequence 48, Appl
28	29	85.3	720	2	US-08-920-919A-48	Sequence 48, Appl
29	28	82.4	106	2	US-09-513-999C-4918	Sequence 4918, Ap
30	28	82.4	126	2	US-09-732-210-261	Sequence 261, Appl
31	28	82.4	144	2	US-08-961-083-44	Sequence 44, Appl
32	28	82.4	144	2	US-09-536-784-44	Sequence 44, Appl
33	28	82.4	144	2	US-09-765-271-44	Sequence 44, Appl
34	28	82.4	144	2	US-09-765-272A-44	Sequence 44, Appl
35	28	82.4	150	2	US-09-248-796A-14820	Sequence 14820, A
36	28	82.4	160	2	US-09-270-767-35707	Sequence 35707, A
37	28	82.4	160	2	US-09-270-767-50924	Sequence 50924, A
38	28	82.4	172	2	US-09-583-110-3217	Sequence 3217, Ap
39	28	82.4	201	2	US-09-107-433-5126	Sequence 5126, Ap
40	28	82.4	221	2	US-09-248-796A-15073	Sequence 15073, A
41	28	82.4	233	2	US-09-328-352-6789	Sequence 6789, Ap
42	28	82.4	251	2	US-09-540-824-9	Sequence 9, Appli
43	28	82.4	252	2	US-09-540-824-1	Sequence 1, Appli
44	28	82.4	281	2	US-09-134-000C-3980	Sequence 3980, Ap
45	28	82.4	292	2	US-09-993-777-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1

US-09-583-110-4900

; Sequence 4900, Application US/09583110

; Patent No. 6699703

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al.

; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus

; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics

; FILE REFERENCE: PATH00-07A

; CURRENT APPLICATION NUMBER: US/09/583,110

; CURRENT FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/107,433

; PRIOR FILING DATE: 1998-06-30

; PRIOR APPLICATION NUMBER: US 60/085,131

; PRIOR FILING DATE: 1998-05-12

; PRIOR APPLICATION NUMBER: US 60/051,553

; PRIOR FILING DATE: 1997-07-02

; NUMBER OF SEQ ID NOS: 5322

; SEQ ID NO 4900

; LENGTH: 126

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

US-09-583-110-4900

Query Match 100.0%; Score 34; DB 2; Length 126;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEYVDN 6

|||||

Db 81 DEYVDN 86

Search completed: December 2, 2005, 09:33:51
Job time : 18.1429 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2005, 09:28:17 ; Search time 94.2857 Seconds
(without alignments)
26.589 Million cell updates/sec

Title: US-10-789-494B-6
Perfect score: 34
Sequence: 1 DEYVDN 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34	100.0	6	5	US-10-789-494B-6	Sequence 6, Appli
2	34	100.0	6	5	US-10-789-494B-66	Sequence 66, Appl
3	34	100.0	120	5	US-10-789-494B-23	Sequence 23, Appl
4	34	100.0	126	5	US-10-617-320-3513	Sequence 3513, Ap
5	34	100.0	148	4	US-10-424-599-274345	Sequence 274345,
6	34	100.0	383	5	US-10-470-048B-576	Sequence 576, App

7	34	100.0	384	4	US-10-724-972A-5636	Sequence 5636, Ap
8	34	100.0	2368	3	US-09-815-242-5635	Sequence 5635, Ap
9	34	100.0	2368	3	US-09-815-242-12389	Sequence 12389, A
10	33	97.1	260	4	US-10-335-977-9215	Sequence 9215, Ap
11	31	91.2	275	4	US-10-438-784-7	Sequence 7, Appli
12	31	91.2	346	5	US-10-450-763-59569	Sequence 59569, A
13	31	91.2	513	5	US-10-732-923-23575	Sequence 23575, A
14	31	91.2	554	3	US-09-895-860-4	Sequence 4, Appli
15	31	91.2	554	4	US-10-377-072-4	Sequence 4, Appli
16	31	91.2	554	4	US-10-377-072-4	Sequence 4, Appli
17	31	91.2	567	4	US-10-369-493-5937	Sequence 5937, Ap
18	31	91.2	582	5	US-10-622-088-104	Sequence 104, App
19	31	91.2	597	4	US-10-369-493-5936	Sequence 5936, Ap
20	31	91.2	704	4	US-10-369-493-5935	Sequence 5935, Ap
21	31	91.2	713	4	US-10-369-493-708	Sequence 708, App
22	31	91.2	983	4	US-10-425-114-38702	Sequence 38702, A
23	31	91.2	1268	4	US-10-438-784-3	Sequence 3, Appli
24	31	91.2	1501	5	US-10-732-923-1749	Sequence 1749, Ap
25	30	88.2	109	4	US-10-424-599-226283	Sequence 226283,
26	30	88.2	143	4	US-10-425-115-311287	Sequence 311287,
27	30	88.2	226	5	US-10-501-282-3888	Sequence 3888, Ap
28	30	88.2	347	4	US-10-767-701-44070	Sequence 44070, A
29	30	88.2	356	4	US-10-156-761-9027	Sequence 9027, Ap
30	30	88.2	428	5	US-10-732-923-10684	Sequence 10684, A
31	30	88.2	468	4	US-10-282-122A-73992	Sequence 73992, A
32	30	88.2	468	4	US-10-282-122A-74751	Sequence 74751, A
33	30	88.2	468	4	US-10-425-114-59773	Sequence 59773, A
34	30	88.2	471	3	US-09-815-242-13277	Sequence 13277, A
35	30	88.2	471	5	US-10-472-928-2368	Sequence 2368, Ap
36	30	88.2	471	5	US-10-617-320-4791	Sequence 4791, Ap
37	30	88.2	495	5	US-10-501-282-3890	Sequence 3890, Ap
38	30	88.2	514	5	US-10-501-282-3892	Sequence 3892, Ap
39	30	88.2	620	4	US-10-424-599-167056	Sequence 167056,
40	30	88.2	621	5	US-10-732-923-18383	Sequence 18383, A
41	30	88.2	647	5	US-10-501-282-3894	Sequence 3894, Ap
42	30	88.2	747	4	US-10-425-115-237165	Sequence 237165,
43	30	88.2	963	5	US-10-732-923-10686	Sequence 10686, A
44	30	88.2	1022	4	US-10-437-963-176734	Sequence 176734,
45	30	88.2	1195	4	US-10-437-963-174724	Sequence 174724,

ALIGNMENTS

RESULT 1

US-10-789-494B-6

; Sequence 6, Application US/10789494B

; Publication No. US20050143296A1

; GENERAL INFORMATION:

; APPLICANT: TSUBOUCHI, Kozo

; APPLICANT: YAMADA, Hiromi

; TITLE OF INVENTION: EXTRACTION AND UTILIZATION OF CELL

; TITLE OF INVENTION: GROWTH-PROMOTING PEPTIDES FROM SILK PROTEIN

; FILE REFERENCE: OPS 635

; CURRENT APPLICATION NUMBER: US/10/789,494B

; CURRENT FILING DATE: 2004-02-27

; PRIOR APPLICATION NUMBER: JP 2003-55048

; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 6
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Antheraea yamamai
US-10-789-494B-6

Query Match 100.0%; Score 34; DB 5; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEYVDN 6
|||
Db 1 DEYVDN 6

RESULT 4

US-10-617-320-3513

; Sequence 3513, Application US/10617320
; Publication No. US20050136404A1
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
FOR DIAGNOSTICS AND
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/617,320
; FILING DATE: 10-Jul-2003
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3513:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...126
; SEQUENCE DESCRIPTION: SEQ ID NO: 3513:
US-10-617-320-3513

Query Match 100.0%; Score 34; DB 5; Length 126;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEYVDN 6
|||
Db 81 DEYVDN 86

Search completed: December 2, 2005, 09:55:40
Job time : 95.2857 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2005, 09:33:57 ; Search time 5.14286 Seconds
(without alignments)
5.586 Million cell updates/sec

Title: US-10-789-494B-6
Perfect score: 34
Sequence: 1 DEYVDN 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 26661 seqs, 4788334 residues

Total number of hits satisfying chosen parameters: 26661

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*

1: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34	100.0	383	6	US-10-793-626-2426	Sequence 2426, Ap
2	29	85.3	1448	6	US-10-485-517-212	Sequence 212, App
3	28	82.4	338	6	US-10-793-626-1756	Sequence 1756, Ap
4	28	82.4	344	7	US-11-083-800-10	Sequence 10, Appl
5	28	82.4	529	6	US-10-821-234-1520	Sequence 1520, Ap
6	28	82.4	1158	7	US-11-075-646-6	Sequence 6, Appli
7	28	82.4	1161	7	US-11-075-646-8	Sequence 8, Appli
8	27	79.4	341	6	US-10-793-626-226	Sequence 226, App
9	27	79.4	532	6	US-10-793-626-546	Sequence 546, App
10	27	79.4	586	6	US-10-131-826A-46	Sequence 46, Appl
11	27	79.4	835	6	US-10-501-039-4	Sequence 4, Appli
12	26	76.5	1304	6	US-10-821-234-1648	Sequence 1648, Ap
13	25	73.5	18	6	US-10-981-873-14	Sequence 14, Appl
14	25	73.5	20	6	US-10-981-873-5	Sequence 5, Appli
15	25	73.5	25	6	US-10-981-873-75	Sequence 75, Appl
16	25	73.5	77	6	US-10-821-234-1132	Sequence 1132, Ap
17	25	73.5	107	6	US-10-467-657-2628	Sequence 2628, Ap
18	25	73.5	182	6	US-10-467-657-4	Sequence 4, Appli
19	25	73.5	182	6	US-10-467-657-3898	Sequence 3898, Ap
20	25	73.5	225	6	US-10-467-657-4472	Sequence 4472, Ap
21	25	73.5	443	6	US-10-793-626-1598	Sequence 1598, Ap
22	25	73.5	443	6	US-10-793-626-1860	Sequence 1860, Ap
23	25	73.5	449	6	US-10-485-517-272	Sequence 272, App
24	25	73.5	554	7	US-11-074-176-320	Sequence 320, App
25	25	73.5	570	7	US-11-074-176-68	Sequence 68, Appl
26	25	73.5	620	7	US-11-055-822-460	Sequence 460, App
27	25	73.5	620	7	US-11-055-822-702	Sequence 702, App
28	25	73.5	709	7	US-11-074-176-158	Sequence 158, App
29	25	73.5	751	6	US-10-821-234-1007	Sequence 1007, Ap
30	25	73.5	895	6	US-10-485-517-129	Sequence 129, App
31	25	73.5	1013	7	US-11-077-550-18	Sequence 18, Appl
32	24	70.6	70	6	US-10-467-657-9208	Sequence 9208, Ap
33	24	70.6	86	6	US-10-467-657-4536	Sequence 4536, Ap
34	24	70.6	229	6	US-10-793-626-1854	Sequence 1854, Ap
35	24	70.6	230	7	US-11-080-628-24	Sequence 24, Appl
36	24	70.6	248	6	US-10-793-626-464	Sequence 464, App
37	24	70.6	266	6	US-10-793-626-2066	Sequence 2066, Ap
38	24	70.6	437	6	US-10-793-626-2960	Sequence 2960, Ap
39	24	70.6	466	6	US-10-467-657-2360	Sequence 2360, Ap

40	24	70.6	532	7	US-11-152-747-2	Sequence 2, Appli
41	24	70.6	546	6	US-10-821-234-902	Sequence 902, App
42	24	70.6	605	6	US-10-689-742-140	Sequence 140, App
43	24	70.6	626	6	US-10-467-657-6426	Sequence 6426, Ap
44	24	70.6	626	6	US-10-467-657-7618	Sequence 7618, Ap
45	24	70.6	661	6	US-10-793-626-274	Sequence 274, App

ALIGNMENTS

RESULT 1

US-10-793-626-2426

```
; Sequence 2426, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2426
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
```

US-10-793-626-2426

```
Query Match          100.0%; Score 34; DB 6; Length 383;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches      6; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
```

```
Qy      1 DEYVDN 6
        |||||
Db      235 DEYVDN 240
```

RESULT 13

US-10-981-873-14

```
; Sequence 14, Application US/10981873
; Publication No. US20050250680A1
; GENERAL INFORMATION:
; APPLICANT: Walensky, Loren D.
; APPLICANT: Korsmeyer, Stanley J.
; APPLICANT: Verdine, Gregory
; TITLE OF INVENTION: STABILIZED ALPHA HELICAL PEPTIDES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 00530-124001
; CURRENT APPLICATION NUMBER: US/10/981,873
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US 60/517,848
```

; PRIOR FILING DATE: 2003-11-05
; PRIOR APPLICATION NUMBER: US 60/591,548
; PRIOR FILING DATE: 2004-07-27
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Naturally occurring peptide
US-10-981-873-14

Query Match 73.5%; Score 25; DB 6; Length 18;
Best Local Similarity 66.7%; Pred. No. 5.4;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEYVDN 6
||:|:
Db 11 DEFVDS 16

RESULT 14

US-10-981-873-5
; Sequence 5, Application US/10981873
; Publication No. US20050250680A1
; GENERAL INFORMATION:
; APPLICANT: Walensky, Loren D.
; APPLICANT: Korsmeyer, Stanley J.
; APPLICANT: Verdine, Gregory
; TITLE OF INVENTION: STABILIZED ALPHA HELICAL PEPTIDES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 00530-124001
; CURRENT APPLICATION NUMBER: US/10/981,873
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US 60/517,848
; PRIOR FILING DATE: 2003-11-05
; PRIOR APPLICATION NUMBER: US 60/591,548
; PRIOR FILING DATE: 2004-07-27
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Naturally occurring peptide
US-10-981-873-5

Query Match 73.5%; Score 25; DB 6; Length 20;
Best Local Similarity 66.7%; Pred. No. 6;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEYVDN 6
||:|:
Db 14 DEFVDS 19

Search completed: December 2, 2005, 09:56:16
Job time : 6.14286 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2005, 09:38:38 ; Search time 20.1429 Seconds
(without alignments)
28.660 Million cell updates/sec

Title: US-10-789-494B-6
Perfect score: 34
Sequence: 1 DEYVDN 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	34	100.0	568	2	F71614	chromatinic RING f
2	34	100.0	2639	2	T31328	fibroin - Chinese
3	33	97.1	218	2	H71866	hypothetical prote
4	33	97.1	714	2	AF0826	lysine decarboxyla
5	31	91.2	282	2	AF0332	hypothetical prote
6	31	91.2	324	2	S57649	concanavalin B pre
7	31	91.2	328	2	C72386	hypothetical prote
8	31	91.2	441	2	T28411	ORF MSV250 hypothe

9	31	91.2	454	2	G75105	hypothetical prote
10	31	91.2	513	2	T14864	probable monosacch
11	31	91.2	554	1	S34607	carboxylesterase (
12	31	91.2	567	2	T33400	protein kinase C h
13	31	91.2	581	1	RGNVBV	trans-activating t
14	31	91.2	582	1	RGNVE2	trans-activating t
15	31	91.2	582	2	E72868	early gene transac
16	31	91.2	582	2	A49626	transregulatory pr
17	31	91.2	597	2	T33399	protein kinase C h
18	31	91.2	636	2	F72867	probable early gen
19	31	91.2	704	1	S60117	protein kinase C (
20	31	91.2	704	2	F86146	hypothetical prote
21	31	91.2	713	2	D85503	lysine decarboxyla
22	31	91.2	713	2	D90652	lysine decarboxyla
23	31	91.2	713	2	B64743	lysine decarboxyla
24	31	91.2	1244	2	T19068	hypothetical prote
25	30	88.2	340	2	T27389	hypothetical prote
26	30	88.2	353	1	WMNV49	40.9K protein - Au
27	30	88.2	353	2	C44221	orf3 protein - Aut
28	30	88.2	353	2	B72852	AcOrf-18 protein -
29	30	88.2	356	2	T41764	AcMNPV orf18 - Bom
30	30	88.2	428	2	T06464	protein kinase (EC
31	30	88.2	468	1	GLSOPL	6-phospho-beta-gal
32	30	88.2	468	2	D95137	6-phospho-beta-gal
33	30	88.2	468	2	D98005	6-phospho-beta-gal
34	30	88.2	477	2	T50551	1-aminocyclopropan
35	30	88.2	737	2	C84232	kinase anchor prot
36	30	88.2	1368	2	T18371	probable glutamate
37	29	85.3	105	2	T19842	hypothetical prote
38	29	85.3	142	2	A44777	profilin spCoell -
39	29	85.3	209	2	D59091	hypothetical prote
40	29	85.3	219	2	AG1940	hypothetical prote
41	29	85.3	250	2	A71268	probable tRNA (gua
42	29	85.3	261	2	T33624	hypothetical prote
43	29	85.3	263	2	E69445	conserved hypothet
44	29	85.3	313	2	T28312	ORF MSV151 probabl
45	29	85.3	354	2	T32246	hypothetical prote

ALIGNMENTS

RESULT 1

F71614

chromatinic RING finger DRING protein homolog PFB0440c - malaria parasite
(Plasmodium falciparum)

C;Species: Plasmodium falciparum

C;Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004

C;Accession: F71614

R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.;
Koonin, E.V.; Shallom, S.; Mason, T.; Yu, K.; Fujii, C.; Pederson, J.; Shen, K.;
Jing, J.; Aston, C.; Lai, Z.; Schwartz, D.C.; Perte, M.; Salzberg, S.; Zhou,
L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.; Fraser, C.M.; Adams,
M.D.; Venter, J.C.; Hoffman, S.L.

Science 282, 1126-1132, 1998

A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium
falciparum.

A;Reference number: A71600; MUID:99021743; PMID:9804551
A;Accession: F71614
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-568 <GAR>
A;Cross-references: UNIPROT:O96182; UNIPARC:UPI000017B5F2; GB:AE001395;
GB:AE001362; NID:g3845184; PIDN:AAC71877.1; PID:g3845185; TIGR:PFB0440c
A;Experimental source: clone 3D7
C;Genetics:
A;Gene: PFB0440c
F;210-260/Domain: RING finger homology <RRN>

Query Match 100.0%; Score 34; DB 2; Length 568;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEYVDN 6
|||||
Db 494 DEYVDN 499

Search completed: December 2, 2005, 09:57:12
Job time : 22.1429 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2005, 09:24:01 ; Search time 125.143 Seconds
(without alignments)
33.827 Million cell updates/sec

Title: US-10-789-494B-6
Perfect score: 34
Sequence: 1 DEYVDN 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	34	100.0		30	2	Q9NBV8_ANTPE	Q9nbv8 antheraea p
2	34	100.0		215	2	Q51E07_ENTHI	Q51e07 entamoeba h
3	34	100.0		383	2	Q5HQ66_STAEQ	Q5hq66 staphylococ
4	34	100.0		383	2	Q8CT10_STAEP	Q8ct10 staphylococ
5	34	100.0		393	2	Q7SDH6_NEUCR	Q7sdh6 neurospora
6	34	100.0		587	2	O96182_PLAF7	O96182 plasmodium
7	34	100.0		1294	2	Q7R6A6_GIALA	Q7r6a6 giardia lam
8	34	100.0		1338	2	Q8IKP8_PLAF7	Q8ikp8 plasmodium
9	34	100.0		1629	2	Q9U0K9_PLAF7	Q9u0k9 plasmodium
10	34	100.0		2639	2	O76786_ANTPE	O76786 antheraea p
11	34	100.0		2655	2	Q964F4_ANTYA	Q964f4 antheraea y
12	33	97.1		84	2	Q5BSE3_SCHJA	Q5bse3 schistosoma
13	33	97.1		204	2	Q4HM18_CAMLA	Q4hm18 campylobact
14	33	97.1		209	2	Q4Y647_PLACH	Q4y647 plasmodium
15	33	97.1		218	2	Q9ZKH7_HELPJ	Q9zkh7 helicobacte
16	33	97.1		290	2	Q7MJF5_VIBVY	Q7mjf5 vibrio vuln
17	33	97.1		314	2	Q8V3H4_SWPV	Q8v3h4 swinepox vi
18	33	97.1		319	2	Q4ZBJ3_9VIRU	Q4zbj3 bacterioph
19	33	97.1		326	2	Q4ZB44_9CAUD	Q4zb44 bacterioph
20	33	97.1		326	2	Q4ZAW9_9CAUD	Q4zaw9 bacterioph
21	33	97.1		358	2	Q5VJ17_AERHY	Q5vj17 aeromonas h
22	33	97.1		714	1	DCLY_SALTI	P0al21 salmonella
23	33	97.1		714	1	DCLY_SALTY	P0al20 salmonella
24	33	97.1		714	2	Q57LF2_SALCH	Q57lf2 salmonella
25	33	97.1		714	2	Q5PIH8_SALPA	Q5pih8 salmonella
26	33	97.1		859	2	Q675L4_PICAB	Q675l4 picea abies
27	33	97.1		965	2	Q4YTG1_PLABE	Q4ytg1 plasmodium
28	33	97.1		967	2	O77305_PLABE	O77305 plasmodium
29	33	97.1		967	2	Q8WP96_PLABE	Q8wp96 plasmodium
30	33	97.1		999	2	Q7RP55_PLAYO	Q7rp55 plasmodium
31	33	97.1		1000	2	Q4I585_GIBZE	Q4i585 gibberella
32	33	97.1		1245	2	Q9U0H6_PLAF7	Q9u0h6 plasmodium
33	33	97.1		1360	2	Q55BM1_DICDI	Q55bm1 dictyosteli
34	33	97.1		1677	2	Q54N52_DICDI	Q54n52 dictyosteli
35	33	97.1		1869	2	Q4YUJ6_PLABE	Q4yuj6 plasmodium
36	33	97.1		5174	2	Q7RTB6_PLAYO	Q7rtb6 plasmodium
37	33	97.1		5251	2	Q8IID4_PLAF7	Q8iid4 plasmodium
38	31	91.2		83	2	Q50V28_ENTHI	Q50v28 entamoeba h
39	31	91.2		133	2	Q50T06_ENTHI	Q50t06 entamoeba h
40	31	91.2		153	2	Q25232_LUCCU	Q25232 lucilia cup
41	31	91.2		172	2	Q9SR34_ARATH	Q9sr34 arabidopsis
42	31	91.2		202	2	Q529E9_MAGGR	Q529e9 magnaporthe
43	31	91.2		205	2	Q4I9V2_GIBZE	Q4i9v2 gibberella
44	31	91.2		263	2	Q8VRR5_9DELT	Q8vrr5 desulfohalo
45	31	91.2		264	2	Q5ZQS6_9DELT	Q5zqs6 desulfohalo

ALIGNMENTS

RESULT 1

Q9NBV8_ANTPE

ID Q9NBV8_ANTPE PRELIMINARY; PRT; 30 AA.
AC Q9NBV8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Fibroin (Fragment).
OS *Antheraea pernyi* (Chinese oak silk moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC Saturniidae; Saturniinae; Saturniini; *Antheraea*.
OX NCBI_TaxID=7119;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Li W., Fan Q., An L.;
RT "Characterization of 5' flanking region for fibroin gene of Chinese
RT Oak Silkworm, *Antheraea pernyi*.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF242774; AAF78030.1; -; Genomic_DNA.
FT NON_TER 30 30
SQ SEQUENCE 30 AA; 3508 MW; F4A680D0F25BD0C4 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEYVDN 6
| | | | |
Db 25 DEYVDN 30

RESULT 2

Q51E07_ENTHI

ID Q51E07_ENTHI PRELIMINARY; PRT; 215 AA.
AC Q51E07;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=12.t00024;
OS *Entamoeba histolytica* HM-1:IMSS.
OC Eukaryota; Entamoebidae; *Entamoeba*.
OX NCBI_TaxID=294381;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HM-1:IMSS;
RX PubMed=15729342; DOI=10.1038/nature03291;
RA Loftus B., Anderson I., Davies R., Alsmark U.C., Samuelson J.,
RA Amedeo P., Roncaglia P., Berriman M., Hirt R.P., Mann B.J., Nozaki T.,
RA Suh B., Pop M., Duchene M., Ackers J., Tannich E., Leippe M.,
RA Hofer M., Bruchhaus I., Willhoeft U., Bhattacharya A.,
RA Chillingworth T., Churcher C., Hance Z., Harris B., Harris D.,
RA Jagels K., Moule S., Mungall K., Ormond D., Squares R., Whitehead S.,
RA Quail M.A., Rabbinowitsch E., Norbertczak H., Price C., Wang Z.,
RA Guillen N., Gilchrist C., Stroup S.E., Bhattacharya S., Lohia A.,

RA Foster P.G., Sicheritz-Ponten T., Weber C., Singh U., Mukherjee C.,
RA El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrell B.,
RA Fraser C.M., Hall N.;
RT "The genome of the protist parasite *Entamoeba histolytica*."
RL Nature 433:865-868(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAFB01000063; EAL51083.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 215 AA; 25532 MW; 51F7F0A900DB894B CRC64;

Query Match 100.0%; Score 34; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEYVDN 6
|||
Db 40 DEYVDN 45

Search completed: December 2, 2005, 09:33:11
Job time : 128.143 secs